Б. Р. ЛЕВИН

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ СТАТИСТИЧЕСКОЙ РАДИОТЕХНИКИ

Б. Р. ЛЕВИН

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ СТАТИСТИЧЕСКОЙ РАДИОТЕХНИКИ

КНИГА ВТОРАЯ



ИЗДАТЕЛЬСТВО «СОВЕТСКОЕ РАДИО»
Москва—1968

УДК 621.396 : ~~519.27~~

Б. Р. ЛЕВИН

Теоретические основы статистической радиотехники. М. Изд-во «Советское радио», 1968, стр. 504, т. 35000 экз.; ц. 1 р. 96 к.

Вторая книга двухтомной монографии по теоретическим основам статистической радиотехники посвящена вопросам оптимального статистического синтеза. Она содержит минимально необходимые сведения из теории решений и теории оценок параметров распределений случайных величин. Подробно освещаются основные положения статистики случайных процессов (ортогональные разложения, функционалы правдоподобия, проверка гипотез, оценки параметров, корреляционных функций и энергетических спектров). Рассматриваются вопросы линейной и нелинейной фильтрации. Излагается теория оптимального синтеза устройств обнаружения и различения сигналов на фоне шума, а также устройств, предназначенных для получения оценок параметров сигналов. Учитывается специфика радиосигналов как узкополосных процессов. Даны элементы теории классификации с обучением. Включены новые результаты, полученные в последние годы.

Книга может служить учебным пособием при изучении статистической радиотехники и дисциплин, примыкающих к ней. Как и первая книга, она рассчитана на научных работников, инженеров, преподавателей, аспирантов и студентов старших курсов радиотехнических вузов.

Табл. 5, рис. 32, библ. назв. 160.

3-4-1
2-68

ПРЕДИСЛОВИЕ

Методы статистической радиотехники, используемые для анализа и синтеза устройств и систем различного назначения, продолжают широким фронтом охватывать области радиотехнических приложений. Однако бурное развитие новых методов и подходов, столь характерное в наше время для многих научных дисциплин, приводит к тому, что зачастую бо́льшая часть потока информации перехлестывает плотину человеческих возможностей и безвозвратно утекает. Оказалось, например, что минимально необходимый интервал времени для написания и издания книги, в котором могут быть обобщены и систематизированы результаты, относящиеся к теоретическим основам статистической радиотехники, соизмерим с периодом, за который абсолютное количество публикуемых работ по этой тематике удваивается.

План задуманной около пяти лет тому назад двухтомной монографии по теоретическим основам статистической радиотехники, реализация которого завершается выходом этой второй книги, не подвергался существенной корректировке. Поэтому многие интересные выводы, полученные математиками и радистами в самые последние годы, не нашли должного отражения как в первой книге*), посвященной методам анализа, так и в представляемой, в которой излагаются методы, характерные для решения задач синтеза. Ограничения, связанные с объемом этой книги и сроком ее издания, явились одной из причин того, что такие актуальные (хотя и не завершенные) разделы, как непараметрические методы, адаптивные системы и стохастические дифференциальные уравнения, либо очень сжато освещены в ней, либо вовсе опущены. Не рассматриваются также современные методы отыскания экстремумов функ-

---

*) Б. Р. Левин. Теоретические основы статистической радиотехники. Книга первая. Изд-во «Советское радио», 1966.

ционалов, составляющих предмет математического программирования. Основное внимание уделено параметрической теории синтеза с полной априорной информацией. Эта часть теории статистического синтеза отличается наибольшей завершенностью и становится уже в известной мере классической.

Книга соответствует второй части аспирантского курса, который читался автором в течение ряда лет. Хотя основными потребителями будут, как и первой книги, аспиранты, научные работники, инженеры-разработчики высокой квалификации и математики, работающие в промышленности, возможно, она найдет место в ряду средств привлечения «неофитов» в лоно статистической радиотехники.

Обе наши книги, объединенные одним и тем же названием, представляют единое целое. Однако те, кто знаком с вероятностными методами анализа, могут пользоваться этой книгой независимо от первой, преодолевая некоторые неудобства, связанные с ссылками на первую книгу, а также с символикой, которая, к сожалению, еще не унифицирована.

Каждый, кому приходилось выпускать в свет печатное слово, знает насколько бывают важны предварительные обсуждения с рецензентами или с коллегами, любезно согласившимися затратить часть своего времени (которого, увы, всегда нехватает) на ознакомление с рукописью. Постоянными помощниками на всех этапах подготовки к изданию этой книги были аспиранты, прослушавшие курс моих лекций. Приношу искреннюю благодарность всем, кому я обязан за внимание к моей работе, в особенности доктору технических наук И. А. Большакову, доктору физико-математических наук Б. С. Флейшману, кандидатам технических наук В. И. Асатуряну, Я. А. Фомину, Ю. Б. Черняку, аспирантам Л. С. Розову и Ю. С. Шинакову. В той или иной степени каждый из них способствовал улучшению изложения, но за оставшиеся незамеченными недостатки повинен исключительно автор этой книги.

Пользуюсь случаем выразить признательность авторам рецензий на мою первую книгу, появившихся в некоторых советских и зарубежных журналах, а также всем, кто прислал мне полезные замечания и пожелания.

Наконец (но не в меньшей степени), я благодарен коллективу сотрудников издательства «Советское радио», который не только помогал мне, но и активно стимулировал работу вот уже на протяжении полутора десятка лет.

Москва, июнь 1967 г.

ВВЕДЕНИЕ

Анализ и синтез. При разработке систем передачи информации (связи, радиолокации, управления и др.) и отдельных элементов этих систем постоянно возникают задачи, которые могут быть классифицированы как задачи двух основных видов: анализа и синтеза. В наиболее общем виде указанные задачи могут быть сформулированы следующим образом. Задача анализа (прямая): заданы характеристики системы (элемента) и процесса, действующего на входе; необходимо найти характеристики процесса на выходе системы (элемента). Задача синтеза (обратная): заданы характеристика процесса на входе и желаемая характеристика процесса на выходе; необходимо найти такую систему, которая преобразовывала бы процесс с заданной характеристикой в процесс с желаемой характеристикой.

В статистической радиотехнике объектом исследований являются различные сигналы и помехи, представляющие реализации *случайных процессов* (и лишь в частных случаях — детерминированных). Задачи анализа при этом состоят в определении необходимых вероятностных характеристик процессов на выходе линейных и нелинейных систем при условии, что построение и характеристики систем заданы и дано вероятностное описание процессов на входе (с той или иной подробностью, необходимой для решения конкретной задачи). Примером анализа является вычисление отношения мощности сигнала к мощности шума на выходе системы, иногда используемое в качестве показателя помехозащищенности. Для решения задач анализа используются математические методы теории вероятностей и теории случайных процессов. Этим методам в их приложении к задачам вероятностного анализа была посвящена наша первая книга [3].

Статистическая формулировка задачи оптимального синтеза состоит в определении *наилучшего* (в некотором смысле) образа действий, позволяющего по наблюдаемой реализа-

ции входного воздействия принять решение о представляющих интерес характеристиках входного воздействия, как случайного процесса (совокупности реализаций). Иначе говоря, необходимо отыскать *оптимальную* систему (алгоритм обработки наблюдаемого входного воздействия), выход которой представлял бы решение или число, характеризующее неизвестные свойства наблюдаемого явления. Выход системы, представлен ли он в форме решения или в числовой форме, является функционалом от наблюдаемой реализации. Для определения вероятностных характеристик этого функционала привлекаются методы, используемые для решения задач анализа.

Следует особенно подчеркнуть, что статистический синтез как построение оптимальной системы или алгоритма обработки связан прежде всего с определением понятия *оптимизации*. Это понятие включает как критерий качества наших действий, так и выдвигаемые нами дополнительные условия, ограничивающие *класс допустимых систем*. Более или менее общий характер оптимизации зависит от вида и числа ограничений, налагаемых на допустимый класс систем. Обоснование выбора критерия качества и ограничений на систему выходит за рамки излагаемой ниже теории синтеза и базируется на человеческом опыте, здравом смысле и интуиции — арсенале средств, принимаемых на вооружение *эвристикой*.

Для решения задач синтеза используются методы математической статистики и теории решений, идейные основы которых были заложены трудами А. Вальда, А. Н. Колмогорова, Дж. Неймана, Н. В. Смирнова, Р. Фишера. Одной из самых первых работ по статистическому синтезу применительно к системам передачи информации была защищавшаяся в 1947 г. докторская диссертация В. А. Котельникова [2], предвосхитившая многие другие (в частности широко известные в настоящее время работы Ф. М. Вудворда [1] и Д. Миддлтона [4]).

Излагаемые в книге методы оптимального статистического синтеза дают в руки разработчиков систем более универсальный инструмент, чем тот, которым они располагали в прежние времена, когда изучение процессов основывалось на более грубых энергетических характеристиках (например, на вездесущем отношении сигнал/шум), которые не отражали особенностей, связанных с выработкой и принятием решений.

Не следует, однако, забывать, что методы статистического синтеза представляют лишь следующую ступень познания, на которой, конечно, не удается полностью освободиться от приближенного рассмотрения объективно существующих явлений и уйти от неизбежных компромиссов, продиктованных выбором критериев качества и необходимостью каким-то образом преодолеть трудности, связанные с отсутствием достаточных априорных данных, с «математическими тупиками», а также со сложностью реализации оптимальных алгоритмов.

Различные критерии, приспособленные к располагаемым априорным данным, приводят, вообще говоря, и к различающимся оптимальным алгоритмам обработки результатов наблюдения, хотя некоторые алгоритмы оказываются иногда оптимальными для ряда критериев. Таковы, например, линейные алгоритмы (согласованные фильтры), оптимальные для обнаружения и различения детерминированных сигналов в аддитивных случайных помехах, имеющих нормальные законы распределения вероятностей.

Обзор содержания. Эта наша вторая книга по теоретическим основам статистической радиотехники, посвященная вопросам статистического синтеза, по своей структуре аналогична первой книге [3].

Первые две главы (первое звено) содержат минимально необходимые сведения из ставших уже классическими теории решений (статистики случайных событий) и теории оценок параметров распределений случайных величин (статистики случайных величин). Изложение этих сведений строго подчинено последующему рассмотрению прикладных вопросов.

Две другие главы (второе звено) представляют математическую основу для решения очень многих задач статистического синтеза, основанного не только на дискретных выборках, но и на *непрерывных реализациях* случайного процесса. Подробно освещаются фундаментальные положения статистики случайных процессов, связанные с ортогональными разложениями случайных процессов, функционалами отношения правдоподобия, проверкой гипотез о случайном процессе, оценками параметров распределения случайного процесса, оценками корреляционных функций и энергетических спектров. Рассматриваются вопросы линейной и нелинейной фильтрации случайных процессов.

Остальная часть книги (третье звено) иллюстрирует специальные приложения теории статистического синтеза для оптимального построения устройств обнаружения и выделения сигналов на фоне шума. Здесь учитывается специфика *радиосигналов* как *узкополосных* процессов. В седьмую главу включены некоторые из основных положений современной теории *классификации* сигналов при неполных априорных данных (распознавания образов).

Небольшое число задач дополняет каждую главу некоторыми представляющими интерес результатами. Самостоятельное решение задач, возможно, укрепит в читателе уверенность в том, что ему под силу статистическое исследование в избранной им области прикладных наук.

Библиография. Как и в первой книге (по указанным там мотивам), мы ограничились в конце каждой главы ссылками в основном на монографическую литературу и на некоторые публикации последних лет, имеющие непосредственное отношение к изложению. Исключение составляет седьмая глава, в которой освещаются вопросы, обсуждавшиеся до сих пор главным образом на страницах периодических изданий. Поэтому список литературы к этой главе более обширный, чем к остальным.

Наконец, следует обратить внимание на вышедшие уже после нашей книги [3] монографии С. М. Рытова [6] и В. И. Тихонова [7]. Все три книги близки по названиям и по целям, однако в отборе материала и в выборе дорог, по которым каждый из трех авторов отправился для достижения этих целей, несомненно, отразилась специфика их научных интересов, методологии и вкусов. Сжатые «Очерки теории связи» Д. Миддлтона [5], крупного американского специалиста в области статистической радиотехники, могут быть хорошим дополнением по вопросам синтеза к книгам упомянутых советских ученых. Последнее замечание в равной мере относится и к предлагаемой книге.

Наша аудитория. Обе книги рассчитаны, прежде всего, на работающих в радиотехнической промышленности инженеров-исследователей и математиков-прикладников. Вопросы синтеза, как и вопросы анализа, являются неотъемлемой частью аспирантского курса статистической радиотехники. Настоящая книга, возникшая вместе с курсом лекций в Московском электротехническом институте связи, может служить пособием для аспирантов, для которых статистическая радиотехника — профилирующий предмет. Ею могут

частично воспользоваться и студенты при изучении курса радиотехнических цепей и сигналов и спецкурсов теории радиосистем.

Опыт работы в промышленности и высшей школе свидетельствует о том, что из года в год неуклонно повышается математический уровень подготовки как молодых инженеров, так и специалистов, уже работающих в научно-исследовательских организациях. Поэтому, хотя мы и не привлекаем дополнительных разделов современной математики (таких, как интеграл Лебега, теория меры и др.) и рассчитываем на читателя, знакомого только с вузовскими курсами математического анализа и основ радиотехники, в некоторых доказательствах и выводах опущены промежуточные выкладки, а ряд утверждений не разъясняется достаточно подробно.

При желании читатель сможет воспроизвести эти подробности самостоятельно, хотя они и не всегда необходимы для понимания существа излагаемой теории. Другое дело — практическое использование теоретических результатов в конкретной практической задаче, когда конечной целью исследователя может быть определение блок-схемы устройства с характеристиками в окончательной числовой (или графической) форме. Здесь часто придется углубляться в математические дебри, в которых, однако, следует искать плоды от деревьев иных пород: не теории вероятностей и статистики, а главным образом линейных интегральных уравнений и линейной алгебры. Мы не злоупотребляли матричной формой записи при использовании многомерных распределений, которая всегда экономит бумагу, но не всегда — мозги.

Во всех случаях нам пришлось жертвовать математической строгостью, когда для ее соблюдения пришлось бы расширить математическую базу или уклониться от основной линии изложения. На первый план поставлена прикладная грань математики, представляющей, по меткому выражению А. Н. Колмогорова на Международном конгрессе математиков в Москве, один из механизмов, при помощи которого человек управляет природой и самим собой.

ЛИТЕРАТУРА

1. Вудворд Ф. М. Теория вероятностей и теория информации с применением в радиолокации. Пер. с англ., под ред. Г. С. Горелика. Изд-во «Советское радио», 1955.
2. Котельников В. А. Теория потенциальной помехоустойчивости. Госэнергоиздат, 1956.
3. Левин Б. Р. Теоретические основы статистической радиотехники. Книга первая. Изд-во «Советское радио», 1966.
4. Миддлтон Д. Введение в статистическую теорию связи. Пер. с англ., под ред. Б. Р. Левина, т. II, ч. IV. Изд-во «Советское радио», 1962.
5. Миддлтон Д. Очерки теории связи. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1966.
6. Рытов С. М. Введение в статистическую радиофизику. Изд-во «Наука», 1966.
7. Тихонов В. И. Статистическая радиотехника. Изд-во «Советское радио», 1966.

ГЛАВА ПЕРВАЯ

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

1.1. ПРОБЛЕМА ВЫБОРА РЕШЕНИЯ

1.1.1. Исходные данные и формулировка проблемы. Рассмотрим совокупность *возможных состояний* $s_0, \ldots, s_m$ (явлений природы, причин появления событий и т. п.), которые представляют полную группу случайных событий, и пусть $p_0, \ldots, p_m$ $(\sum_{k=0}^{m} p_k = 1)$ — априорное распределение вероятностей этих состояний. Рассмотрим далее совокупность результатов наблюдений $x_1, \ldots, x_n$ (выборочных значений), зависящих от того, какое из упомянутых состояний в действительности имеет место, и пусть $W_n(x_1, \ldots, x_n|s_k)$ — условное распределение выборочных значений, соответствующее состоянию s_k, $k = 0,1, \ldots, m$.

Имеются: набор решений $\gamma_0, \ldots, \gamma_m$ относительно истинности состояний, *правила* $\delta(\gamma_i|x_1, \ldots, x_n)$ выбора решения, приписывающие каждому возможному результату наблюдений $x_1, \ldots, x_n$ одно из решений γ_i, $i = 0,1, \ldots, m$, а также *функция потерь* $\Pi(s_j, \gamma_k)$, учитывающая последствия выбора решения, и *критерий качества* $f\{\Pi\}$ выбора решения, связанный с функцией потерь.

Проблема, которая будет подробно рассмотрена ниже и которая лежит в основе многих задач техники и естествознания, может быть в общем виде сформулирована следующим образом: при заданных распределениях p_k, $W_n(x_1, \ldots, x_n|s_k)$, наборе решений γ_i, функции потерь Π и критерии качества f требуется определить наилучшее в смысле принятого критерия *правило* δ использования результатов наблюдений $x_1, \ldots, x_n$ для *выбора решения*. Это правило является разновидностью статистического вывода, получаемого по результатам наблюдений, о неизвестных сторонах изучаемого явления, точнее говоря, о принятой математической модели в условиях

неполной информации относительно характеристик этой модели.

Прежде чем детализировать введенные понятия и постановку проблемы выбора решения, проиллюстрируем их на очень простом примере.

1.1.2. Пример: передача сообщений при наличии помех. Система связи состоит из источника, производящего два сообщения: «да» и «нет», которые преобразуются соответственно в сигналы «1» и «0», канала с помехами, которые могут преобразовывать сигнал «1» в сигнал «0» или «0» в «1», и приемного устройства, воспроизводящего сигналы в неискаженном или искаженном виде. Оператор, наблюдающий сигнал на выходе приемного устройства, должен расшифровать передачу, т. е. определить по принятому сигналу переданное сообщение.

Если бы сигналы всегда передавались без искажений, то можно было бы по принятому сигналу дать безошибочный ответ на вопрос, какое сообщение было послано. Например, по сигналу «1» всегда узнавалось бы сообщение «да». Иначе говоря, апостериорная вероятность сообщения «да» при условии, что принят сигнал «1», равнялась бы единице. Вследствие искажений помехами принятый сигнал не всегда будет достоверно указывать на то, какое сообщение было передано, т. е. будут случаи, когда принимается сигнал «1» при передаче сообщения «нет», а сигнал «0» — при передаче сообщения «да». Возникает необходимость дать оператору заранее правило поведения в указанной неопределенной ситуации, не полагаясь на его интуицию и субъективные суждения.

Рассмотренный простейший пример содержит все элементы приведенной выше постановки проблемы. Сообщения «нет» и «да» представляют два взаимонесовместимых состояния передатчика s_0 и s_1. Априорные вероятности этих состояний $p_0 = P\{\text{«нет»}\}$, $p_1 = P\{\text{«да»}\} = 1 - p_0$ определяют статистическую структуру источника сообщений, т. е. указывают, какой процент в длинной последовательности сообщений составляют сообщения «нет» и какой процент — сообщения «да». Возможными результатами наблюдений являются сигналы «1» и «0». Условные вероятности этих сигналов $P\{0 \mid s_0\} = 1 - P\{1 \mid s_0\}$, $P\{1 \mid s_1\} = 1 - P\{0 \mid s_1\}$ определяются вероятностными свойствами помех в канале. Величины $P\{0 \mid s_0\}$, $P\{1 \mid s_1\}$ представляют вероятности того, что сигналы «0» и «1» не искажаются

помехами, а $P\{1 \mid s_0\}$, $P\{0 \mid s_1\}$ — вероятности искажений двух видов: перехода «0» в «1» и «1» в «0». Набор решений в этом случае состоит из γ_0 (передано сообщение «нет») и γ_1 (передано сообщение «да»), а правило решения предписывает оператору, какое из этих двух решений он должен выбирать, когда наблюдает сигнал «1» или сигнал «0». Функция потерь в рассматриваемом примере должна учитывать последствия ошибочных решений оператора и назначать «плату» $\Pi_{01} > 0$ за ошибку первого рода (принятие решения, что было передано сообщение «да», когда в действительности передавалось «нет») и «плату» $\Pi_{10} > 0$ за ошибку второго рода (принятие решения, что было передано сообщение «нет», когда в действительности передавалось «да»). Можно было бы наряду с потерями Π_{01}, Π_{10} ввести величины выигрышей Π_{00}, Π_{11} (отрицательных потерь), приобретаемых от правильных решений или расходов на принятие правильных решений (при условии, что $\Pi_{00} < \Pi_{01}$ и $\Pi_{11} < \Pi_{10}$). Однако иногда ограничиваются введением только потерь из-за ошибочных решений, полагая, что правильным решениям соответствуют нулевые потери.

Критерием качества выбора решения может служить среднее значение потерь из-за ошибочных решений, взвешенное с вероятностями их появления. Таким образом, согласно этому критерию выбирается из двух возможных правил выбора решения то, для которого величина среднего значения потерь меньше. Подсчитаем величины средних потерь для двух правил выбора каждого из возможных решений.

Одно правило δ_0 может быть сформулировано так: наблюдаешь сигнал «0»— принимай решение γ_0 (и, следовательно, когда наблюдаешь сигнал «1»— принимай решение γ_1). В этом случае вероятности ошибочных решений равны

$$P\{\text{ошибка 1-го рода}\} = P\{\gamma_1 \mid s_0\} = p_0 P\{1 \mid s_0\},$$
$$P\{\text{ошибка 2-го рода}\} = P\{\gamma_0 \mid s_1\} = p_1 P\{0 \mid s_1\}$$

и среднее значение R_0 потерь

$$R_0 = \Pi_{01} p_0 P\{1 \mid s_0\} + \Pi_{10} p_1 P\{0 \mid s_1\} \qquad (1.1)$$

Другое правило δ_1 формулируется так: наблюдаешь сигнал «0»— принимай решение γ_1 (и, следовательно, когда наблюдаешь сигнал «1»— принимай решение γ_0). В этом

случае вероятности ошибочных решений равны

$$P\{\text{ошибка 1-го рода}\} = P\{\gamma_0 \mid s_0\} = p_0 P\{0 \mid s_0\},$$
$$P\{\text{ошибка 2-го рода}\} = P\{\gamma_1 \mid s_1\} = p_1 P\{1 \mid s_1\}$$

и среднее значение R_1 потерь

$$R_1 = \Pi_{01} p_0 P\{0 \mid s_0\} + \Pi_{10} p_1 P\{1 \mid s_1\}. \qquad (1.2)$$

Принятый критерий качества отдает предпочтение правилу δ_0, если $R_0 < R_1$, т. е. когда

$$R_0 < \frac{1}{2}(R_0 + R_1). \qquad (1.3)$$

Так как $R_0 + R_1 = \Pi_{01} p_0 + \Pi_{10} p_1$, то из (1.1) и (1.3) приходим к следующему условию, при выполнении которого принимается правило δ_0:

$$\Pi_{01} p_0 P\{1 \mid s_0\} + \Pi_{10} p_1 P\{0 \mid s_1\} < \frac{1}{2}(\Pi_{01} p_0 + \Pi_{10} p_1). \qquad (1.4)$$

Формула (1.4) помимо условных вероятностей ошибок, определяемых вероятностными характеристиками помех в канале, содержит априорные вероятности сообщений и величины потерь. Определение или назначение величин p_0, Π_{01}, Π_{10} в конкретных ситуациях может представлять значительные трудности. Это обстоятельство является слабым местом и в общей постановке проблемы, указанной в § 1.1.1. Когда нет никаких оснований для того, чтобы ошибку первого рода считать более или менее существенной, чем ошибку второго рода, то полагают потери Π_{01} и Π_{10} одинаковыми, и тогда величины средних потерь просто пропорциональны вероятности ошибки любого рода. Критерий наименьших средних потерь в этом случае переходит в критерий наименьшей частоты ошибок. Когда ничего «неизвестно» о статистической структуре источника сообщений, то остается предположить, что сообщения «да» и «нет» передаются с равными вероятностями, т. е. $p_0 = p_1 = \frac{1}{2}$.

Если $\Pi_{01} = \Pi_{10}$, $p_0 = p_1 = \frac{1}{2}$, то условие (1.4) становится особенно простым:

$$P\{1 \mid s_0\} + P\{0 \mid s_1\} < 1$$

или

$$P\{1 \mid s_0\} < P\{1 \mid s_1\}. \qquad (1.5)$$

Условие (1.5) означает, что вероятность искажения сигнала «0» меньше вероятности правильного воспроизведения сигнала «1» (иначе, вероятность появления ложного сигнала «1» меньше вероятности неподавления истинного сигнала «1»). Средние потери в этом случае $R_0 < 0{,}5$, в то время как для правила δ_1 при условии (1.5) средние потери $R_1 > 0{,}5$ *).

1.1.3. Простые и сложные гипотезы. Теперь рассмотрим несколько подробнее основные понятия, введенные в § 1.1.1 в связи с общей постановкой проблемы выбора решения. Иногда эта проблема формулируется в терминах теории *проверки статистических гипотез.* Выбор решения тогда состоит в принятии или отклонении гипотезы относительно возможных состояний изучаемого явления по результатам наблюдений. Так, например, при обнаружении сигнала на фоне шума результат наблюдений выходного эффекта приемного устройства может относиться только к шуму (состояние s_0) или к смеси сигнала и шума (состояние s_1). Выбор решения состоит в принятии или отклонении гипотезы H_0 о том, что наблюдавшийся эффект относится только к шуму. Противоположная гипотеза H_1 о том, что наблюдавшийся эффект относится к смеси сигнала и шума, является альтернативой.

Класс гипотез называется *простым*, если он содержит одну-единственную гипотезу, и *сложным*, если число гипотез не меньше двух. В указанном выше примере проверяемая гипотеза и альтернатива являются простыми. В задаче обнаружения одного из совокупности сигналов на фоне шума (состояния $s_1, \ldots, s_m$) класс альтернатив является сложным. Сложной гипотезе может соответствовать не только конечное число состояний, но и бесконечное: счетное (дискретное) множество или континуум состояний. Возможные состояния принято представлять точками многомерного пространства состояний. Если множество возможных состояний образует континуум, то априорное распределение вероятностей состояний характеризуется плот-

---

*) Заметим, что возможными правилами решения в рассматриваемом примере могли бы быть также: δ_3— всегда принимай решение γ_0, или δ_4 — всегда принимай решение γ_1. Величины средних потерь для этих решений равны соответственно $R_3 = p_1\Pi_{10}$, $R_4 = p_0\Pi_{01}$. Если $\Pi_{10} = \Pi_{10} = 1$ и $p_0 = p_1 = {}^1/_2$, то любое из этих правил хуже, чем δ_0, так как при этом $R_3 = R_4 = {}^1/_2$, но лучше, чем δ_1, так как $R_1 > 0{,}5$.

ностью вероятности $w_1(s)$, определенной на этом множестве.

1.1.4. Выборка. Как отмечалось в первой книге, с каждым случайным экспериментом может быть связана некоторая случайная величина ξ, возможные значения которой представляют результаты наблюдений, фиксируемые после проведения случайного эксперимента. Таким образом, результат последовательности n случайных экспериментов представляется n возможными значениями $x_1, \ldots, x_n$ случайной величины ξ. Каждое из этих значений называется выборочным, а их совокупность — *выборкой* *). Число n выборочных значений, или число элементов выборки, принято называть *размером* (или объемом) выборки. Пусть распределение вероятностей случайной величины ξ характеризуется интегральной функцией $F_1(x)$ или плотностью $w_1(x)$. Тогда говорят, что выборка $x_1, \ldots, x_n$ получена из распределения $F_1(x)$ или соответственно из $w_1(x)$. Закон распределения может зависеть от состояния s_k изучаемого явления. Для того чтобы подчеркнуть зависимость выборки от указанного состояния s_k, будем обозначать функцию распределения, из которого она получена, символом $F_1(x \mid s_k)$ или $w_1(x \mid s_k)$.

Каждая выборка, т. е. совокупность n чисел, представляет точку в n-мерном пространстве, или n-мерный вектор $\mathbf{X} = (x_1, \ldots, x_n)$. Если рассматривать (хотя бы мысленно) все возможные последовательности экспериментов, то совокупность всевозможных выборок заполняет некоторую область указанного n-мерного пространства. Эта область называется *пространством выборок*. Плотность вероятности выборки из распределения $w_1(x \mid s)$ с независимыми элементами $x_1, \ldots, x_n$ равна

$$W_n(x_1, \ldots, x_n \mid s) = \prod_{k=1}^{n} w_1(x_k \mid s), \tag{1.6}$$

так как в этом случае совместная плотность вероятности выборочных значений равна произведению плотностей вероятностей элементов выборки.

*) Во многих руководствах по математической статистике совокупность возможных значений случайной величины ξ принято называть *генеральной совокупностью*. Каждый случайный эксперимент представляет *выбор* некоторого значения из генеральной совокупности.

Для выборки из дискретного распределения $p_{j\,|\,s} = P\{x = x_j\,|\,s\}$, $j = 1, 2, \ldots, n$,

$$P\{x = x_1, \ldots, x = x_n\,|\,s\} = \prod_{k=1}^{n} p_{k\,|\,s}. \qquad (1.6')$$

Совместное распределение выборочных значений называется *функцией правдоподобия* выборки. Функция правдоподобия, заданная в пространстве выборок, указывает, как часто появляются точки из некоторой области пространства выборок, и определяет плотность вероятности того, что появляется какая-то точка пространства выборок.

Итак, принимается точка зрения на выборку как на совокупность случайных величин или как на многомерную случайную величину (случайный вектор), характеризующуюся некоторым распределением вероятностей — функцией правдоподобия *).

Обобщая понятие выборки, можно предположить, что результатами наблюдений являются значения не одной случайной величины, а совокупности зависимых случайных величин, связанных n-мерной функцией распределения $W_n(x_1, \ldots, x_n\,|\,s)$, вид которой определяется состоянием s изучаемого явления.

До сих пор предполагалось, что в процессе извлечения выборочных значений $x_1, \ldots, x_n$ состояние s_k изучаемого явления не изменялось. В примере с обнаружением это означало, что наблюдаемый эффект связан либо только с шумом, либо со смесью сигнала и шума. Однако могут быть и такие ситуации, когда состояние изучаемого явления изменяется один или несколько раз в процессе извлечения выборочных значений. Например, сигнал в шумах появляется в пределах заданного интервала наблюдения. Тогда выборка оказывается *разнородной*, отдельные ее части принадлежат разным распределениям $w_1(x\,|\,s_k)$, $w_1(x\,|\,s_j)$ и т. д.

Наконец, обратим внимание на то, что в рассмотренной выше постановке «сырьем» для статистических выводов

---

*) Заметим, что в отличие от первой книги, где случайные величины и аргументы соответствующих им функций распределения обозначались разными символами (греческими и латинскими), здесь элементы выборки и аргументы соответствующей функции правдоподобия обозначаются одинаковыми символами, что, однако, не должно служить поводом для отождествления этих величин.

служит выборка *заданного* размера n. Иногда оказывается целесообразнее не планировать заранее объем экспериментов, а определять его в процессе наблюдений, принимая решение о прекращении или продолжении наблюдений *последовательно* после каждого эксперимента, подвергая статистическому анализу накопленные к данному моменту выборочные значения.

1.1.5. Набор решений и правило выбора решения. Набор решений $\gamma_0, \gamma_1, \ldots, \gamma_m$ представляет собой ряд логических утверждений о том, какая из гипотез относительно состояний $s_0, s_1, \ldots, s_m$ изучаемого явления истинна. Правило выбора решения δ устанавливает соответствие между набором решений и возможными результатами наблюдений, т. е. пространством выборок*). Это означает, что пространство выборок G должно быть разделено на $m+1$ непересекающихся областей $G_0, G_1, \ldots, G_m$, и тогда правило выбора решения устанавливает соответствие между решениями γ_k и областями $G_j(x_1, \ldots, x_n)$. Важно подчеркнуть, что правило выбора решения устанавливается *до* проведения наблюдения.

Правило решения может быть *детерминированным* (или нерандомизированным). При этом данной области $G_k(x_1, \ldots, x_n)$ всегда приписывается определенное решение γ_k, иначе говоря, если наблюдаемая выборка попадает в область G_k, то принимается решение γ_k, т. е. утверждается истинность гипотезы о том, что изучаемое явление находится в состоянии s_k. Правило решения может быть *рандомизированным*. При этом для заданных выборочных значений $x_1, \ldots, x_n$ допускается выбор одного из нескольких решений в соответствии с некоторым распределением вероятностей. Это распределение $P\{\gamma_k \mid (x_1, \ldots, x_n) \in G_j\}$ представляет условные вероятности решений при фиксированной выборке $\left(\sum_{k=0}^{m} P\{\gamma_k \mid (x_1, \ldots, x_n) \in G_j\} = 1, \quad j = 0, 1, \ldots, m\right)$. Для детерминированного правила решения $P\{\gamma_k \mid (x_1, \ldots, x_n) \in G_j\} = 1$ лишь для одного значения $k = j$, и указанная вероятность равна нулю для всех других значений $k \neq j$. В дальнейшем в этой книге будут рассматриваться только детерминированные правила

---

*) В математической литературе само правило выбора решения часто называют *критерием*. Мы используем термин «критерий» только в смысле критерия качества правила выбора решения.

выбора решений. Этим ограничением, однако, не игнорируется роль рандомизированных правил, которые имеют самостоятельное значение и дают иногда возможность упростить математический анализ (см., например, [7]).

1.1.6. Функция потерь и критерий качества выбора решения. Использование любого заранее установленного правила выбора решения в силу случайной природы объекта наблюдения неминуемо связано с возможностью ошибочных решений. Наблюдаемая выборка $x_1, \ldots, x_n$ может оказаться в области G_k, и будет принято решение γ_k, что изучаемое явление находится в состоянии s_k, хотя в действительности указанная выборка связана с другим состоянием s_j, $j \neq k$. Наличие в последовательности решений не только правильных, но и ошибочных является *неизбежной* платой за попытку получить решение в условиях неполной информации.

Последствия ошибочных решений могут быть различны. Аналитически это обстоятельство учитывается путем введения неотрицательной функции потерь, которая предписывает каждому ошибочному решению, т. е. каждой комбинации s_j и γ_k, $k \neq j$, плату $\Pi_{jk} = \Pi(s_j, \gamma_k) > 0$. Наряду с этим могут быть введены также величины выигрышей (отрицательных потерь), приобретаемых при правильных решениях, или расходы, связанные с вынесением правильных решений $\Pi_{jj} = \Pi(s_j, \gamma_j) < \Pi_{jk}$, $k \neq j$. Для заданного состояния s_j средняя величина потерь при использовании определенного правила δ выбора решения (т. е. способа разбиения пространства выборок на области G_k и установления их соответствия набору решений γ_k) в достаточно длинной последовательности экспериментов (результаты которых фиксируются выборками размера n) приближенно равна среднему в пространстве выборок значению (математическому ожиданию) функции потерь

$$r_j = \sum_{k=0}^{m} \Pi_{jk} P\{\gamma_k \mid s_j\} = \sum_{k=0}^{m} \Pi_{jk} P\{(x_1, \ldots, x_n) \in G_k \mid s_j\}, \quad (1.7)$$

где $P\{(x_1, \ldots, x_n) \in G_k \mid s_j\}$ — условная вероятность попадания выборки $(x_1, \ldots, x_n)$ в область G_k, если в действительности имеет место состояние s_j. Условное среднее r_j для данного состояния s_j называется *условной функцией риска*.

Можно было принять условную функцию риска за критерий качества правила выбора решения и считать наилуч-

шим правилом то, которое минимизирует по всем возможным δ величину r_j. Однако оптимальное свойство правила решения зависело бы при этом от заданного состояния s_j. Для другого состояния s_k, $k \neq j$, правило, минимизирующее r_j, возможно, уже не будет реализовывать наименьшее значение r_k.

Усредняя условную функцию риска по всем возможным состояниям s_j, получим

$$R = \sum_{j=0}^{m} p_j r_j = \sum_{j=0}^{m} \sum_{k=0}^{m} p_j \Pi_{jk} P\{(x_1, \ldots, x_n) \in G_k \mid s_j\}, \qquad (1.8)$$

где p_j — априорная вероятность состояния s_j.

Определенное согласно (1.8) среднее значение функции потерь, которое зависит и от правила выбора решения, и от распределения вероятностей состояний, называется *средней функцией риска*. Эта функция может быть принята за критерий качества правила выбора решения. Тогда оптимальным (наилучшим в смысле принятого критерия) правилом выбора решения будет такое, для которого при заданных функции потерь, распределении вероятностей состояний и условных распределениях вероятностей выборок при фиксированных состояниях величина средней функции риска будет наименьшей среди значений этой функции, относящихся к любым иным правилам выбора решения. Оптимальным правилом выбора решения задается один (из многообразия возможных) способ разбиения пространства выборок на непересекающиеся области G_k, $k = 0, 1, \ldots, m$, который обеспечивает при длительном его использовании минимальные средние (по возможным состояниям изучаемого явления) потери.

Оптимальное правило выбора решения, минимизирующее среднюю функцию риска, называется *байесовским решением*, а соответствующее ему минимальное значение средней функции риска — *байесовским риском*.

Излагаемая теория обладает существенным недостатком. Прежде чем воспользоваться ее результатами, необходимо запастись достаточно обширной априорной информацией не только об условных функциях распределения выборочных значений $W_n(x_1, \ldots, x_n \mid s_j)$, которые во многих случаях можно обоснованно задать, но и о функции потерь $\Pi(s_j, \gamma_k)$ и об априорном распределении состояний.

Если априорное распределение состояний неизвестно, то для установления критерия качества выбора решения

имеется возможность использовать лишь условную функцию риска r_j, которая представляет функцию целочисленного аргумента j, указывающего состояние s_j. Пусть r^*_{i1} — максимальное значение этой функции для правила выбора решения δ_1 и r^*_{k2} — ее максимальное значение для правила выбора решения δ_2. Можно считать более благоприятным то из этих двух правил, для которого максимальное значение условной функции риска меньше. Например, если $r^*_{k2} < r^*_{i1}$, то правило δ_2 лучше правила δ_1. Оптимальным правилом будет такое, которому соответствует минимум среди максимальных значений условных функций риска, определенных при любых других правилах. Это правило называют *минимаксным* правилом выбора решения.

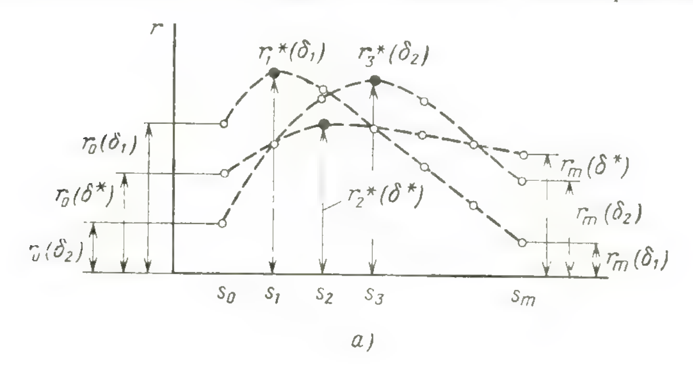
Минимаксное правило выбора решения дает уверенность в том, что потери (в среднем) не будут больше некоторой величины r^*. Хотя во многих случаях использование этого правила разумно, могут быть ситуации, в которых оно будет слишком осторожным, и предпочтительным окажется иное правило решения, для которого максимальное значение условного риска будет больше, чем r^*. Так, например, если в одном из состояний правило выбора решения δ приводит к условному риску, несколько большему, чем r^*, а в остальных состояниях — намного меньшему, чем те, которые соответствуют минимаксному правилу δ^*, то целесообразно отдать предпочтение правилу δ. Указанное обстоятельство иллюстрируется на рис. 1, на котором изображены типичные случаи, когда использование минимаксного правила обоснованно (рис. 1, *а*) и когда минимаксное правило является чересчур осторожным (рис. 1,*б*).

Можно показать (см., например, [14, стр. 91]), что *любое* минимаксное правило выбора решения является *специальным случаем байесовского решения для наименее благоприятного априорного распределения вероятностей* $(p_j)_{\text{мм}}$ состояний s_j, $j = 0, \ldots, m$, для которого минимальный (байесовский) средний риск имеет наибольшую величину по сравнению с величинами среднего риска, вычисленными для байесовского решения при любом другом распределении *). К сожалению, не существует общего метода, при помощи которого можно было бы находить наименее благоприятное априорное распределение $(p_j)_{\text{мм}}$.

---

*) Условия существования такого наименее благоприятного распределения сформулированы в работе [14, § 3.1, теорема 3.14].

Однако доказано (см. [14, § 3.5]), что байесовское правило выбора решения, которому соответствуют одинаковые для всех состояний условные риски $r_j = r$, $j = 0, 1, \ldots, m$, является минимаксным. Это обстоятельство может быть использовано для определения наименее благоприятного



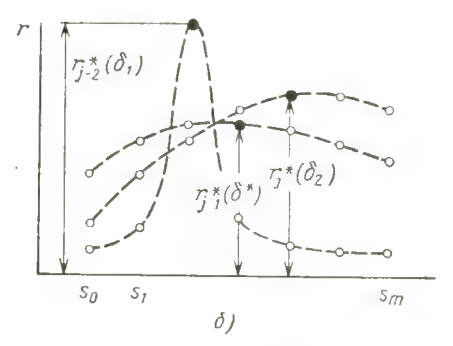


Рис. 1. Использование минимаксного правила:
а — приемлемо; *б* — неприемлемо (правило δ_1 лучше, чем δ^*).

распределения и, следовательно, минимаксного правила. Но следует иметь в виду, что равенство условных рисков при *небайесовских* решениях не приводит к минимаксному правилу. С другой стороны, если равенство условных рисков неосуществимо, то это еще не значит, что минимаксного правила не существует. Для его определения в этом случае нужно искать другие методы.

В условиях, когда известно априорное распределение состояний s_j, но нет каких-либо обоснованных соображений

относительно величин потерь Π_{jk}, может быть использован несколько иной подход к выработке правил выбора решений. Найдем, используя формулу Байеса (см. (1.19) в первой книге), апостериорную вероятность состояния s_j, когда наблюдается выборка $(x_1, \ldots, x_n)$:

$$P\{s_j \mid x_1, \ldots, x_n\} = \frac{p_j W_n(x_1, \ldots, x_n \mid s_j)}{\sum\limits_{k=0}^{m} p_k W_n(x_1, \ldots, x_n \mid s_k)}. \qquad (1.9)$$

Указанные апостериорные вероятности представляют наиболее полную характеристику состояния s_j изучаемого явления, какую можно получить, располагая только выборочными значениями случайной величины, связанной с наблюдениями. Поэтому естественно принять следующий критерий: утверждается истинной та из гипотез относительно состояний s_j, $j = 0, 1, \ldots, m$, для которой апостериорная вероятность (1.9) максимальна. Таким образом, критерием качества рассматриваемого правила выбора решения является *максимум апостериорной вероятности*.

Из этого критерия следует правило разбиения пространства выборок. К области G_k относят те выборки $(x_1, \ldots, x_n)$, для которых при всех $j \neq k$.

$$P\{s_k \mid x_1, \ldots, x_n\} \geqslant P\{s_j \mid x_1, \ldots, x_n\}. \qquad (1.9')$$

Если нет никаких априорных сведений и о распределении вероятностей состояний, и о потерях, то можно воспользоваться критерием *максимального правдоподобия*, согласно которому при наблюдении выборки $(x_1, \ldots, x_n)$ принимается та из гипотез относительно состояний s_j, для которой функция правдоподобия $W_n(x_1, \ldots, x_n \mid s_j)$ больше других функций правдоподобия $W_n(x_1, \ldots, x_n \mid s_k)$, $k = 0, 1, \ldots, j-1, j+1, \ldots, m$. Этот критерий является частным случаем критерия максимальной апостериорной вероятности при условии, что все состояния равновероятны, т. е. $p_j = \frac{1}{m+1}$.

Еще один критерий качества правила выбора решения связан со средним *количеством информации*, содержащимся в используемом правиле δ относительно исследуемого явления s (характеризуемого $m+1$ состояниями $s_0, \ldots, s_m$ и их априорными распределениями $p_0, \ldots, p_m$). По опреде-

лению среднего количества информации [13]

$$I(\delta,\ s)=\sum_{j=0}^{m}\sum_{k=0}^{m}P\{s_j\ \text{и}\ \gamma_k\}\ln\frac{P\{s_j\ \text{и}\ \gamma_k\}}{p_jP\{\gamma_k\}}=$$

$$=H(s)-H(s\,|\,\delta)=H(s)-\sum_{k=0}^{m}P\{\gamma_k\}H(s\,|\,\gamma_k), \qquad (1.10)$$

где

$$H(s)=-\sum_{j=0}^{m}p_j\ln p_j\geqslant 0$$

— *энтропия*, характеризующая априорную неопределенность состояний исследуемого явления, и

$$H(s\,|\,\gamma_k)=-\sum_{j=0}^{m}P\{s_j\,|\,\gamma_k\}\ln P\{s_j\,|\,\gamma_k\}\geqslant 0$$

— *условная энтропия* явления s после того, как принято решение γ_k, причем

$$P\{s_j\,|\,\gamma_k\}=\frac{p_j}{P\{\gamma_k\}}\int\limits_{G_k}\cdots\int W_n(x_1,\ \ldots,\ x_n\,|\,s_j)\,dx_1\ldots dx_n,$$

$$P\{\gamma_k\}=\sum_{j=0}^{m}p_j\int\limits_{G_k}\cdots\int W_n(x_1,\ \ldots,\ x_n\,|\,s_j)\,dx_1\ldots dx_n,$$

$$H(s\,|\,\delta)=\sum_{k=0}^{m}P\{\gamma_k\}H(s\,|\,\gamma_k),\quad 0\leqslant H(s\,|\,\delta)\leqslant H(s).$$

В качестве наилучшего правила δ может быть принято такое, которое доставляет максимум информации $I(\delta,\ s)$, т. е. минимизирует среднее значение $H(s\,|\,\delta)$ условной энтропии (или, как принято говорить в теории информации, *ненадежности*). Иначе говоря, правило выбора решения, удовлетворяющее условию $\max\limits_{\delta} I(\delta,\ s)$, гарантирует наименьшие в среднем потери информации, связанные с процессом принятия решения на основе выборочных значений *).

*) Более полный информационный критерий качества связан с максимизацией *ценности информации* при условии, что количество полученной информации не превосходит заданной величины [2].

1.2. ПРОВЕРКА ПРОСТОЙ ГИПОТЕЗЫ ПРОТИВ ПРОСТОЙ АЛЬТЕРНАТИВЫ

1.2.1. Вероятности правильных и ошибочных решений. Переходим к простейшей задаче излагаемой в этой главе теории — проверке простых гипотез. Ситуация в этом случае такова. Имеется некоторое число наблюденных значений $x_1, \ldots, x_n$ (выборка размера n) и известно, что эти значения принадлежат одному из двух распределений: $W_n(x_1, \ldots, x_n \mid s_0)$ или $W_n(x_1, \ldots, x_n \mid s_1)$, связанных с взаимоисключающими состояниями s_0 и s_1 изучаемого явления. Задача состоит в том, чтобы указать наилучший (в каком-нибудь смысле) алгоритм обработки наблюдаемых данных с целью решить, какому из указанных распределений принадлежит полученная выборка.

Обозначим через H_0 и H_1 — гипотезы о том, что выборочные значения принадлежат распределениям $W_n(x_1, \ldots, x_n \mid s_0)$ и $W_n(x_1, \ldots, x_n \mid s_1)$ соответственно, а через γ_0 и γ_1 — решения, состоящие в принятии или отклонении гипотезы H_0. Гипотеза H_1 является простой альтернативой H_0, и поэтому может рассматриваться только одна гипотеза H_0. Ясно, что отклонение гипотезы H_0 означает принятие гипотезы H_1. Для рассматриваемых здесь *нерандомизированных* процедур проверки гипотезы (см. § 1.1.5) задача состоит в установлении до наблюдений правила, согласно которому каждой выборке $x_1, \ldots, x_n$ приписывалось бы одно из решений γ_0 или γ_1, иначе говоря, в установлении правила, по которому можно было бы принять или отвергнуть гипотезу H_0 на основании данных, накопленных в процессе наблюдения изучаемого явления. Установление указанного правила эквивалентно разделению n-мерного пространства выборок $(x_1, \ldots, x_n)$ на две непересекающиеся области G_0 и G_1. Если данная конкретная выборка попадает в область G_0, то гипотеза H_0 принимается, а если она попадает в область G_1, то она отвергается (т. е. принимается гипотеза H_1). Таким образом,

$$(x_1, \ldots, x_n) \in G_0 \rightarrow \gamma_0,$$

$$(x_1, \ldots, x_n) \in G_1 \rightarrow \gamma_1,$$

где символ включения $\in$ означает принадлежность точки данной области пространства.

Область G_0 принятия гипотезы называют *допустимой*, а область G_1 отклонения гипотезы — *критической*. Уравне-

ние поверхности $D(x_1, \ldots, x_n) = \text{const}$ в n-мерном пространстве, разделяющей указанные области, является аналитическим выражением правила выбора решений.

При использовании любого заранее установленного правила выбора решений наряду с правильными решениями неизбежны (в силу случайной природы выборки) и *ошибочные*. Возможны ошибки двух родов. *Ошибка первого рода* возникает, когда выборка попадает в критическую область G_1, когда изучаемое явление находится в состоянии s_0. Тем самым будет отвергнута гипотеза H_0, хотя в действительности она верна. *Ошибка второго рода* возникает, когда выборка попадает в допустимую область G_0, хотя изучаемое явление находится в состоянии s_1. В результате будет принята ложная гипотеза. Аналогично могут рассматриваться и два вида правильных решений; принятие верной гипотезы (выборка попадает в область G_0, когда имеет место состояние s_0) и отклонение ложной гипотезы (выборка попадает в область G_1, когда имеет место состояние s_1).

Нетрудно написать выражения для условных вероятностей ошибок для заданного состояния изучаемого явления. Условная вероятность α ошибки первого рода равна

$$\begin{aligned}\alpha = P\{\gamma_1 | H_0\} = P\{(x_1, \ldots, x_n) \in G_1 | s_0\} = \\ = \int\limits_{G_1} \ldots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n.\end{aligned} \qquad (1.11)$$

Условная вероятность правильного решения, состоящего в принятии верной гипотезы H_0, дополняет указанную вероятность до единицы, т. е.

$$\begin{aligned}P\{\gamma_0 | H_0\} = P\{(x_1, \ldots, x_n) \in G_0 | s_0\} = \\ = \int\limits_{G_0} \ldots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n = \\ = 1 - \int\limits_{G_1} \ldots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n = 1 - \alpha.\end{aligned} \qquad (1.11')$$

Условная вероятность β ошибки второго рода равна

$$\begin{aligned}\beta = P\{\gamma_0 | H_1\} = P\{(x_1, \ldots, x_n) \in G_0 | s_1\} = \\ = \int\limits_{G_0} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n.\end{aligned} \qquad (1.12)$$

Условная вероятность правильного решения, состоящего в отклонении ложной гипотезы, дополняет β до единицы, так как

$$P\{\gamma_1 | H_1\} = P\{(x_1, \ldots, x_n) \in G_1 | s_1\} = \int\limits_{G_1} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n =$$

$$= 1 - \int\limits_{G_0} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n = 1 - \beta. \quad (1.12')$$

Вероятность α ошибки первого рода (т. е. вероятность отвергнуть правильную гипотезу H_0) называют иногда *уровнем значимости*, а вероятность $1 - \beta$ отвергнуть ложную гипотезу — *мощностью* правила выбора решений *).

Если известно, что априорные вероятности состояний s_0 и s_1 равны q и $p = 1 - q$ соответственно, то, используя формулы (1.11) — (1.12′), можно найти априорные вероятности принятия решений γ_0 и γ_1:

$$P\{\gamma_0\} = qP\{\gamma_0 | H_0\} + pP\{\gamma_0 | H_1\} = q(1 - \alpha) + p\beta, \quad (1.13)$$

$$P\{\gamma_1\} = pP\{\gamma_1 | H_1\} + qP\{\gamma_1 | H_0\} = p(1 - \beta) + q\alpha, \quad (1.14)$$

которые определяют частоты появления отдельных решений в длинной последовательности принятия решений. В формулах (1.13) и (1.14) первые слагаемые равны априорным вероятностям правильных решений, а вторые — априорным вероятностям ошибок.

Для заданного размера выборки невозможно одновременно сделать сколь угодно малыми вероятности ошибок и первого, и второго рода. Например, чтобы уменьшить уровень значимости, нужно уменьшить критическую область G_1.

---

*) Иногда в пространстве выборок вводят решающую функцию $\Phi(X)$, которая для нерандомизированного правила выбора решения равна

$$\Phi(X) = \begin{cases} 1, & X \in G_1, \\ 0, & X \in G_0, \end{cases}$$

т. е. является своеобразным счетчиком попадания выборки $X = (x_1, \ldots, x_n)$ в критическую область. Формулы (1.11) и (1.12′) могут быть при помощи решающей функции $\Phi(X)$ записаны в виде условных средних

$$\alpha = m_1\{\Phi(X) | H_0\}, \quad 1 - \beta = m_1\{\Phi(X) | H_1\}.$$

Для рандомизированного правила $\Phi(X)$ представляет некоторую интегральную функцию распределения.

При этом, конечно, увеличится допустимая область и понизится чувствительность правила решения в отношении ошибок второго рода. Поэтому для того, чтобы сформулировать то или иное правило выбора решений, необходимо выработать какие-то разумные подходы. Путь к таким подходам указывают критерии качества, некоторые из них были представлены в общем виде в § 1.1.6. Рассмотрим теперь эти критерии применительно к задаче проверки простых гипотез.

1.2.2. Байесовское решение. Введем, прежде всего, функцию потерь, которая предписывает каждой из четырех комбинаций γ_0 и H_0, γ_0 и H_1, γ_1 и H_0, γ_1 и H_1 соответствующую плату Π_{jk}, $j=0,1$, $k=0,1$. Последние величины удобно представить в виде платежной матрицы

$$\Pi=\begin{pmatrix}\Pi_{00} & \Pi_{01}\\ \Pi_{10} & \Pi_{11}\end{pmatrix},\quad \Pi_{01}>\Pi_{00},\quad \Pi_{10}>\Pi_{11}, \tag{1.15}$$

в которой строки соответствуют гипотезам H_0 и H_1, а столбцы — решениям γ_0 и γ_1. По главной диагонали расположены платы за правильные решения *), а по побочной — платы (потери) за ошибочные решения. Среднее значение потерь, взвешенное с вероятностями их появления, или *средний риск*) равно

$$R=qr_0+pr_1, \tag{1.16}$$

где

$$r_0=\Pi_{00}P\{\gamma_0\mid H_0\}+\Pi_{01}P\{\gamma_1\mid H_0\}=\Pi_{00}(1-\alpha)+\Pi_{01}\alpha, \tag{1.17}$$

$$r_1=\Pi_{10}P\{\gamma_0\mid H_1\}+\Pi_{11}P\{\gamma_1\mid H_1\}=\Pi_{10}\beta+\Pi_{11}(1-\beta) \tag{1.17'}$$

— условные риски, соответствующие состояниям s_0 и s_1. Подставляя (1.17) и (1.17′) в (1.16), после простых преобразований получим

$$R=q\Pi_{00}+p\Pi_{10}+q(\Pi_{01}-\Pi_{00})\alpha-p(\Pi_{10}-\Pi_{11})(1-\beta). \tag{1.18}$$

Примем в качестве критерия для определения правила выбора решений достижение минимальной величины среднего риска R. В рассматриваемом случае средний риск R

*) Если $\Pi_{00}<0$, $\Pi_{11}<0$, то эти величины можно трактовать как выигрыши при принятии правильных решений.

зависит от критической области G_1 через величины α и $1-\beta$. Подставляя выражения этих величин из (1.11) и (1.12′) в (1.18), находим

$$R = q\Pi_{00} + p\Pi_{10} - \int_{G_1}\!\!\dots\!\int [p(\Pi_{10}-\Pi_{11}) W_n(x_1, \dots, x_n \mid s_1) - \\ - q(\Pi_{01}-\Pi_{00}) W_n(x_1, \dots, x_n \mid s_0)]\, dx_1 \dots dx_n. \quad (1.19)$$

Так как $q\Pi_{00} + p\Pi_{10}$ — постоянная величина, то минимальное значение среднего риска R будет получено в том случае, когда подынтегральная функция в (1.19) будет неотрицательной *), т. е. когда в критическую область G_1 пространства выборок (где отвергается гипотеза H_0) включаются только те точки, для которых

$$p(\Pi_{10}-\Pi_{11}) W_n(x_1, \dots, x_n \mid s_1) \geqslant \\ \geqslant q(\Pi_{01}-\Pi_{00}) W_n(x_1, \dots, x_n \mid s_0)$$

или

$$\frac{pW_n(x_1, \dots, x_n \mid s_1)}{qW_n(x_1, \dots, x_n \mid s_0)} \geqslant \frac{\Pi_{01}-\Pi_{00}}{\Pi_{10}-\Pi_{11}}. \quad (1.20)$$

Таким образом, оптимальное правило, основанное на критерии минимального среднего риска, или *байесовское решение*, может быть сформулировано следующим образом: принимается решение γ_1 (отвергается гипотеза H_0), если для наблюдаемой выборки выполняется неравенство (1.20), и принимается решение γ_0 (утверждается справедливость гипотезы H_0), если выполняется неравенство, противоположное (1.20).

Уравнение поверхности, разделяющей в этом случае критическую и допустимую области пространства выборок, имеет вид

$$D(x_1, \dots, x_n) = \frac{pW_n(x_1, \dots, x_n \mid s_1)}{qW_n(x_1, \dots, x_n \mid s_0)} = \frac{\Pi_{01}-\Pi_{00}}{\Pi_{10}-\Pi_{11}}. \quad (1.21)$$

---

*) Заметим, что так как для любого подмножества G_1' множества G_1 ($G_1' \subset G_1$) справедливо неравенство

$$\int_{G_1'}\!\!\dots\!\int f(x_1, \dots, x_n)\, dx_1 \dots dx_n \leqslant \int_{G_1}\!\!\dots\!\int f(x_1, \dots, x_n)\, dx_1 \dots dx_n$$

при $f(x_1, \dots, x_n) \geqslant 0$, то интеграл в (1.19) достигает максимума тогда и только тогда, когда в область интегрирования G_1 включаются все точки пространства выборок, для которых подынтегральная функция неотрицательна.

Левую часть этого уравнения $\frac{pW_n(x_1, \ldots, x_n \mid s_1)}{qW_n(x_1, \ldots, x_n \mid s_0)}$ называют *обобщенным отношением правдоподобия* *). Процедура проверки простой гипотезы H_0 сводится к вычислению обобщенного отношения правдоподобия и сравнению его с постоянным порогом c^*, который для байесовского правила равен **)

$$c^* = \frac{\Pi_{01} - \Pi_{00}}{\Pi_{10} - \Pi_{11}}. \tag{1.22}$$

Если обозначить через G_1^* критическую область, определяемую неравенством (1.20), а через α^* и $1 - \beta^*$ — уровень значимости и мощность, соответствующие G_1^*, то минимальная величина среднего риска R^*, или *байесовский риск*, может быть представлена согласно (1.18) в следующем виде:

$$R^* = q\Pi_{00} + p\Pi_{10} + q(\Pi_{01} - \Pi_{00})\alpha^* - p(\Pi_{10} - \Pi_{11})(1 - \beta^*). \tag{1.23}$$

При определении условных вероятностей ошибок α^* и β^* можно обойти трудности, связанные с вычислением кратных интегралов (1.11) и (1.12), и свести вычисления к однократному интегрированию. Функция

$$l(x_1, \ldots, x_n) = \frac{W_n(x_1, \ldots, x_n \mid s_1)}{W_n(x_1, \ldots, x_n \mid s_0)}, \tag{1.24}$$

т. е. *отношение правдоподобия*, представляет неотрицательную случайную величину, получающуюся функциональным преобразованием ***) n случайных величин $x_1, \ldots, x_n$. Обозначим через $W_{10}(y)$ функцию распределения отношения правдоподобия при условии, что выборка принадлежит распределению $W_n(x_1, \ldots, x_n \mid s_0)$, и через $W_{11}(y)$ —

---

*) В отличие от величины $l(x_1, \ldots, x_n) = \frac{W_n(x_1, \ldots, x_n \mid s_1)}{W_n(x_1, \ldots, x_n \mid s_0)}$, которую называют просто *отношением правдоподобия*. (Некоторые авторы предпочитают термин «коэффициент правдоподобия».)

**) Хотя при выводе указанного оптимального правила предполагалось заранее, что рассматриваются только детерминированные правила, полученный результат остается справедливым и в том случае, когда минимизируется средний риск в более широком классе, включающем и рандомизированные правила (см., например, [9, § 19.1.2] или [10, § 2.3]).

***) Функциональное преобразование $z = l(x_1, \ldots, x_n)$ отображает точки n-мерного пространства выборок на действительную ось (одномерную область).

функцию распределения отношения правдоподобия при условии, что выборка принадлежит распределению $W_n(x_1, \ldots, x_n \mid s_1)$. Тогда из (1.11) и (1.12) и определения критической области G_1^* [см. (1.20) и (1.22)] следует:

$$\alpha^* = P\{(x_1, \ldots, x_n) \in G_1^* \mid s_0\} =$$

$$= P\{l(x_1, \ldots, x_n) \geqslant \mu c^* \mid s_0\} = \int\limits_{\mu c^*}^{\infty} W_{10}(y)\, dy = 1 - F_{10}(\mu c^*), \tag{1.25}$$

$$\beta^* = P\{(x_1, \ldots, x_n) \in G_0^* \mid s_1\} =$$

$$= P\{l(x_1, \ldots, x_n) < \mu c^* \mid s_1\} = F_{11}(\mu c^*), \tag{1.26}$$

где F_{10}, F_{11} — интегральные функции распределения отношения правдоподобия при гипотезах H_0 и H_1 соответственно и $\mu = \frac{q}{p}$.

1.2.3. Максимум апостериорной вероятности и максимальное правдоподобие. Определим апостериорные вероятности того, что изучаемое явление находится в состоянии s_0 или в состоянии s_1, когда наблюдается выборка $x_1, \ldots, x_n$. Из (1.9) в рассматриваемом случае находим

$$P\{s_0 \mid x_1, \ldots, x_n\} = \frac{qW_n(x_1, \ldots, x_n \mid s_0)}{qW_n(x_1, \ldots, x_n \mid s_0) + pW_n(x_1, \ldots, x_n \mid s_1)}, \tag{1.27}$$

$$P\{s_1 \mid x_1, \ldots, x_n\} = \frac{pW_n(x_1, \ldots, x_n \mid s_1)}{qW_n(x_1, \ldots, x_n \mid s_0) + pW_n(x_1, \ldots, x_n \mid s_1)}. \tag{1.28}$$

Установим следующее правило выбора решения: после того как получена выборка $x_1, \ldots, x_n$, принимается гипотеза H_0, если $P\{s_0 \mid x_1, \ldots, x_n\} > P\{s_1 \mid x_1, \ldots, x_n\}$ (решение γ_0), и отвергается эта гипотеза, если $P\{s_0 \mid x_1, \ldots, x_n\} \leqslant P\{s_1 \mid x_1, \ldots, x_n\}$ (решение γ_1). Используя (1.27) и (1.28), можно это правило представить в виде: принимается решение γ_1 (отвергается гипотеза H_0), если для наблюдаемой выборки выполняется неравенство

$$\frac{pW_n(x_1, \ldots, x_n \mid s_1)}{qW_n(x_1, \ldots, x_n \mid s_0)} \geqslant 1, \tag{1.29}$$

и принимается решение γ_0 (справедлива гипотеза H_0), если выполняется неравенство, противоположное (1.29).

Таким образом, максимуму апостериорной вероятности соответствует такая критическая область пространства выборок, точки которой удовлетворяют неравенству (1.29). Процедура проверки простой гипотезы H_0 сводится в этом случае к вычислению обобщенного отношения правдоподобия и сравнению его с единицей. Нетрудно заметить, сравнивая (1.29) с (1.20), что рассматриваемое правило выбора решения является просто частным случаем байесовского решения, когда порог $c^* = 1$ [см. (1.22)]. Это соответствует равенству плат за решения γ_0 и γ_1 либо равенству плат за ошибки $\Pi_{10} = \Pi_{01} = \Pi$, если принимается, что $\Pi_{00} = \Pi_{11} = 0$. В последнем случае средний риск оказывается равным [см. (1.18)]

$$R = (q\alpha + p\beta)\Pi, \tag{1.30}$$

т. е. отличается только константой Π от априорной вероятности ошибок любого рода. Следовательно, правило выбора решения по критерию максимальной апостериорной вероятности *минимизирует априорную вероятность ошибок*. Иначе говоря, это правило на протяжении длинной последовательности принятия решений обеспечивает максимальную частоту *правильных решений*.

Если при выработке правила выбора решений нет никаких данных относительно априорных вероятностей состояний s_0 и s_1, то вместо рассмотренного критерия можно применить критерий максимума правдоподобия, согласно которому при наблюдении выборки $x_1, \ldots, x_n$ принимается та из гипотез, которой соответствует большее значение функции правдоподобия выборки. Таким образом, принимается гипотеза H_0, если

$$W_n(x_1, \ldots, x_n \mid s_0) > W_n(x_1, \ldots, x_n \mid s_1)$$

(решение γ_0), и отвергается эта гипотеза, если

$$W_n(x_1, \ldots, x_n \mid s_0) \leqslant W_n(x_1, \ldots, x_n \mid s_1)$$

(решение γ_1). Иначе говоря, принимается решение γ_1, если для наблюдаемой выборки

$$l(x_1, \ldots, x_n) = \frac{W_n(x_1, \ldots, x_n \mid s_1)}{W_n(x_1, \ldots, x_n \mid s_0)} \geqslant 1, \tag{1.31}$$

и принимается решение γ_0, если выполняется неравенство, противоположное (1.31). Процедура проверки простой гипотезы по критерию максимального правдоподобия сводится

к вычислению *отношения правдоподобия* и сравнению его с единицей. Отличие от правила, соответствующего максимуму апостериорной вероятности, состоит лишь в замене обобщенного отношения правдоподобия

$$\frac{pW_n(x_1, \ldots, x_n \mid s_1)}{qW_n(x_1, \ldots, x_n \mid s_0)}$$

отношением правдоподобия

$$\frac{W_n(x_1, \ldots, x_n \mid s_1)}{W_n(x_1, \ldots, x_n \mid s_0)}.$$

Следовательно, правило выбора решения по критерию максимального правдоподобия является частным случаем правила по критерию максимума апостериорной вероятности, когда два возможных состояния изучаемого явления s_0 и s_1 *равновероятны*, т. е. когда $p = q = \frac{1}{2}$.

1.2.4. Критерий Неймана — Пирсона. Другой подход к выработке правила выбора решений при отсутствии априорной информации о потерях и вероятностях состояний указывает критерий Неймана — Пирсона. Согласно этому критерию выбирается такое правило, которое обеспечивает минимально возможную величину β вероятности ошибок второго рода при условии, что вероятность ошибки первого рода не больше заданной величины α. Иначе говоря, правило выбора решения по критерию Неймана — Пирсона имеет *наибольшую мощность* среди всех других правил, для которых *уровень значимости не превосходит* α.

Покажем, что указанному правилу соответствует критическая область G_1 пространства выборок, включающая только те точки, для которых

$$l(x_1, \ldots, x_n) = \frac{W_n(x_1, \ldots, x_n \mid s_1)}{W_n(x_1, \ldots, x_n \mid s_0)} \geqslant c, \tag{1.32}$$

причем порог c выбирается из условия

$$P\{l(x_1, \ldots, x_n) \geqslant c \mid s_0\} = \int_c^\infty W_{10}(y)\,dy = \alpha. \tag{1.33}$$

Пусть T — произвольная область пространства выборок, удовлетворяющая единственному условию

$$P\{(x_1, \ldots, x_n) \in T \mid s_0\} \leqslant \alpha. \tag{1.34}$$

Обозначим через U пересечение областей G_1 и T, через A — ту часть области G_1, которая не пересекается с T, и че-

рез B — ту часть области T, которая не пересекается с G_1 (рис. 2) *). Так как точки области A удовлетворяют неравенству (1.32), то

$$\int_A \cdots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n \geqslant$$
$$\geqslant c \int_A \cdots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n.$$

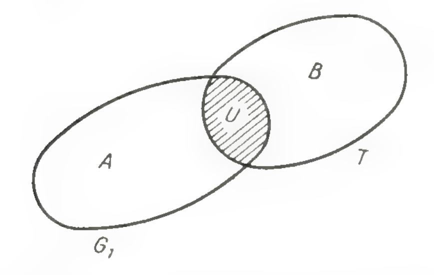


Рис. 2. Области пространства выборок.

Поэтому

$$\int_{G_1} \cdots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n =$$
$$= \int_A \cdots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n +$$
$$+ \int_U \cdots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n \geqslant$$
$$\geqslant c \int_A \cdots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n +$$
$$+ \int_U \cdots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n =$$
$$= c \int_{G_1} \cdots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n -$$
$$- c \int_U \cdots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n +$$
$$+ \int_U \cdots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n. \qquad (1.35)$$

*) Не исключаются случаи, когда $A \equiv G_1$, $B \equiv T$, $U \equiv 0$.

И так как в силу (1.33) и (1.34)

$$\alpha = \int_{G_1} \ldots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n \geqslant$$

$$\geqslant \int_{T} \ldots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n,$$

то из (1.35) следует:

$$\int_{G_1} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n \geqslant$$

$$\geqslant c \int_{T} \ldots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n -$$

$$- c \int_{U} \ldots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n +$$

$$+ \int_{U} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n =$$

$$= c \int_{B} \ldots \int W_n(x_1, \ldots, x_n | s_0)\, dx_1 \ldots dx_n +$$

$$+ \int_{U} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n.$$

Область B не содержит точек области G_1, и поэтому в соответствии с (1.32) в этой области

$$cW_n(x_1, \ldots, x_n | s_0) > W_n(x_1, \ldots, x_n | s_1).$$

Следовательно,

$$\int_{G_1} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n >$$

$$> \int_{B} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n +$$

$$+ \int_{U} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n =$$

$$= \int_{T} \ldots \int W_n(x_1, \ldots, x_n | s_1)\, dx_1 \ldots dx_n, \qquad (1.36)$$

т. е. мощность правила по критерию Неймана — Пирсона [см. (1.26)], равная

$$\int_{G_1}\ldots\int W_n(x_1, \ldots, x_n \mid s_1)\,dx_1\ldots dx_n = 1-\beta = 1-F_{11}(c), \tag{1.37}$$

больше мощности любого другого правила, удовлетворяющего условию (1.34).

Из сравнения (1.32) с (1.20) можно заключить, что правило выбора решений, базирующееся на критерии Неймана — Пирсона, является *частным случаем* байесовского решения, в котором величина μc^* заменяется на c, определяемую согласно (1.33).

Итак, все рассмотренные выше критерии качества приводят к единообразной процедуре принятия решения: по наблюденной выборке $x_1, \ldots, x_n$ вычисляется отношение правдоподобия $l(x_1, \ldots, x_n)$ и принимается или отвергается гипотеза H_0 в зависимости от того, находится ли эта величина ниже или выше некоторого фиксированного порога, устанавливаемого заранее в соответствии с принятым критерием.

1.2.5. Способ вычисления условных вероятностей ошибок. Если $\Phi(y)$ — монотонная функция, то сравнение $l(x_1, \ldots, x_n)$ с порогом c *эквивалентно* сравнению $\Phi[l(x_1, \ldots, x_n)]$ с порогом $\Phi(c)$. Оптимальность процедуры проверки гипотезы не нарушается. Условные вероятности ошибок вычисляются по формулам вида (1.25), (1.26), в которых распределения отношений правдоподобия заменяются распределениями $\Phi(l)$, а в пределы интегрирования вместо порога c подставляется величина $\Phi(c)$.

Когда выборка состоит из независимых элементов, функция правдоподобия выборки представляется произведением их одномерных функций распределения [см. (1.6)]. Такой вид функции правдоподобия подсказывает выбор функции Φ в виде *логарифма* отношения правдоподобия. При этом оказывается возможным представить логарифм многомерного отношения правдоподобия в виде суммы логарифмов одномерных отношений правдоподобия:

$$\ln l(x_1, \ldots, x_n) = \sum_{i=1}^{n} \ln l(x_i) = \sum_{i=1}^{n} \ln\frac{w_1(x_i \mid s_1)}{w_1(x_i \mid s_0)} \tag{1.38}$$

и свести определение функции распределения $\ln l(x_1, \ldots, x_n)$ к определению функции распределения суммы *независимых* одинаково распределенных случайных величин $\ln l(x_i)$. Обозначим через $\Theta_1(v \mid s_0)$, $\Theta_1(v \mid s_1)$ — условные характеристические функции случайной величины $\ln l(x)$ при гипотезах H_0 и H_1 соответственно. Тогда в соответствии с известным в теории вероятностей соотношением [см. (3.115) в первой книге]

$$\Theta_1(v \mid s_0) = \int_{-\infty}^{\infty} w_1(x \mid s_0) \exp\left[iv \ln \frac{w_1(x \mid s_1)}{w_1(x \mid s_0)}\right] dx, \quad (1.39)$$

$$\Theta_1(v \mid s_1) = \int_{-\infty}^{\infty} w_1(x \mid s_1) \exp\left[iv \ln \frac{w_1(x \mid s_1)}{w_1(x \mid s_0)}\right] dx, \quad (1.40)$$

а плотности вероятностей случайной величины $\ln l(x_1, \ldots, x_n)$ равны

$$W_1(z \mid s_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\Theta_1(v \mid s_0)]^n e^{-ivz} dv, \quad (1.41)$$

$$W_1(z \mid s_1) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\Theta_1(v \mid s_1)]^{nz} e^{-ivz} dv. \quad (1.42)$$

Теперь, используя (1.25) и (1.26), можно определить условные вероятности ошибочных решений по формулам

$$\alpha = \int_{\ln c}^{\infty} W_1(z \mid s_0) dz = 1 - F_1(\ln c \mid s_0), \quad (1.43)$$

$$\beta = \int_{-\infty}^{\ln c} W_1(z \mid s_1) dz = F_1(\ln c \mid s_1). \quad (1.44)$$

Здесь $F_1(z \mid s_0)$ и $F_1(z \mid s_1)$ — условные интегральные функции распределения $\ln l(x_1, \ldots, x_n)$.

Если n велико и условные дисперсии $M_2\{\ln l(x) \mid s_i\}$ ограничены, то распределение $\ln l(x_1, \ldots, x_n)$ как суммы большого числа одинаково распределенных независимых случайных величин [см. (1.38)] согласно теореме Ляпунова (см. § 3.4 в первой книге) приближается к нормальному.

В этом случае

$$W_1(z \mid s_0) \sim \frac{1}{\sqrt{2\pi n M_{20}}} \exp\left\{-\frac{(z-nm_{10})^2}{2nM_{20}}\right\}, \qquad (1.45)$$

$$W_1(z \mid s_1) \sim \frac{1}{\sqrt{2\pi n M_{21}}} \exp\left\{-\frac{(z-nm_{11})^2}{2nM_{21}}\right\}, \qquad (1.46)$$

где

$$m_{10} = m_1\{\ln l(x) \mid s_0\}, \ \ M_{20} = M_2\{\ln l(x) \mid s_0\}, \qquad (1.47)$$

$$m_{11} = m_1\{\ln l(x) \mid s_1\}, \ \ M_{21} = M_2\{\ln l(x) \mid s_1\}. \qquad (1.48)$$

Тогда имеют место следующие *асимптотические* соотношения [см. (1.43) и (1.44)]:

$$\alpha \sim 1 - F\left(\frac{\ln c - nm_{10}}{\sqrt{nM_{20}}}\right), \qquad (1.49)$$

$$\beta \sim F\left(\frac{\ln c - nm_{11}}{\sqrt{nM_{21}}}\right), \qquad (1.50)$$

где $F(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$

(интеграл Лапласа, см. § 2.2.2 в первой книге).

Можно показать (см. задачу 1.10), что $m_{10} < 0$ и $m_{11} > 0$. Отсюда следует, что при $n \to \infty$ одновременно $\alpha \to 0$ и $\beta \to 0$ (для критерия Неймана — Пирсона при заданном α стремится к нулю β).

Правило решения, для которого вероятность ошибки асимптотически (при $n \to \infty$) стремится к нулю, называется *состоятельным*.

1.2.6. Минимаксное правило. Исследуем также минимаксное правило выбора решения, воспользовавшись тем обстоятельством, что это правило представляет специальный случай байесовского решения для наименее благоприятного априорного распределения состояний изучаемого явления (см. § 1.1.6). Так как в рассматриваемом случае проверки простых гипотез имеется лишь два возможных состояния s_0 и s_1, то указанное распределение определяется лишь одной вероятностью q (или $p = 1 - q$). Поэтому для того, чтобы найти то значение q, с которым связано наибольшее значение байесовского риска, необходимо определить максимум величины R^* как функции q, используя (1.23) и имея в виду, что эта переменная входит в правую часть (1.23) как непосредственно, так и через α^* и β^* [см. (1.25) и (1.26)].

Дифференцируя правую часть (1.23) по q, приравнивая результат дифференцирования нулю и учитывая, что точки поверхности, разделяющей критическую и допустимую области пространства выборок, удовлетворяют условию (1.21), приходим к трансцендентному уравнению относительно искомого наименее благоприятного значения вероятности q

$$\Pi_{00}[1-\alpha^*(q)]+\Pi_{01}\alpha^*(q)=\Pi_{10}\beta^*(q)+\Pi_{11}[1-\beta^*(q)]. \quad (1.51)$$

Уравнение (1.51) непосредственно следует из (1.17) и (1.17′), так как оно просто выражает равенство условных рисков $r_0=r_1$ для байесовского решения. Это равенство, как указывалось в § 1.1.6, и характеризует минимаксное правило выбора решения. Подставляя (1.25) и (1.26) в (1.51), можно указанное трансцендентное уравнение относительно q $\left(\text{или } \mu=\frac{q}{p}\right)$ привести к виду

$$\frac{\Pi_{00}-\Pi_{11}}{\Pi_{10}-\Pi_{11}}+c^*[1-F_{10}(\mu c^*)]=F_{11}(\mu c^*). \quad (1.51')$$

Решая уравнение (1.51′) и выделяя тот корень уравнения $\mu_{мм}$ (и, следовательно, $q_{мм}$), которому соответствует абсолютный максимум байесовского риска, приходим к следующему *минимаксному правилу* выбора решения: принимается решение γ_1 (отвергается гипотеза H_0), если для наблюдаемой выборки $x_1, \ldots, x_n$ выполняется неравенство

$$\frac{p_{мм}}{q_{мм}}l(x_1, \ldots, x_n)\geqslant c^*, \quad (1.52)$$

и решение γ_0 (утверждается справедливость гипотезы H_0), если выполняется неравенство, противоположное (1.52).

Величина минимаксного риска в соответствии с (1.23), (1.25) и (1.26) может быть вычислена по формуле

$$R_{мм}=q_{мм}\Pi_{00}+p_{мм}\Pi_{10}+q_{мм}(\Pi_{01}-\Pi_{00})\alpha_{мм}-\\-p_{мм}(\Pi_{10}-\Pi_{11})(1-\beta_{мм}), \quad (1.53)$$

где

$$\alpha_{мм}=1-F_{10}(\mu_{мм}c^*),\ \beta_{мм}=F_{11}(\mu_{мм}c^*). \quad (1.54)$$

Разность $R_{мм}-R^*\geqslant 0$ между минимаксным (при неизвестном q) и байесовским риском (при известном q) является платой за отсутствие априорной информации о состояниях изучаемого явления.

Резюмируя сказанное в § 1.2.2—1.2.5, можно отметить, что для всех рассмотренных критериев процедура проверки простой гипотезы сводится к сравнению отношения правдоподобия с порогом *) c. Выражения, определяющие этот порог для различных критериев, сведены в табл. 1.

Таблица 1

Критерий	Порог c
Байесовский	$\frac{q}{p} \frac{\Pi_{01}-\Pi_{00}}{\Pi_{10}-\Pi_{11}}$
Максимум апостериорной вероятности	$\mu = \frac{q}{p}$
Максимум правдоподобия	1
Неймана—Пирсона	Из уравнения $F_{10}(c) = 1-\alpha$
Минимаксный	$\mu_{мм} \frac{\Pi_{01}-\Pi_{00}}{\Pi_{10}-\Pi_{11}}$ $\mu_{мм}$ из (1.51)

1.2.7. Последовательный анализ. Отличительной особенностью всех рассмотренных процедур выбора решения при проверке простой гипотезы была неизменность заранее заданного размера выборки n. Существует другой подход к установлению правила выбора решения, при котором отказываются от постоянного размера выборки, а ограничивают эту величину в процессе эксперимента в зависимости от результатов уже выполненных наблюдений. Вначале наблюдают одно-единственное значение x_1 (извлекают выборку размера $n = 1$) и на основании этого значения по заранее установленному правилу выбирают одно из трех решений:

1) принять гипотезу H_0 (решение γ_0), 2) отвергнуть гипотезу H_0, т. е. принять гипотезу H_1 (решение γ_1), 3) *про-*

*) Заметим, что критическая область, для которой $l(x_1, \ldots, x_n) \geqslant c$, не обязательно *односвязная* (см. задачу 1.5).

должить наблюдения, т. е. отказаться на данном этапе от принятия одного из решений, γ_0 или γ_1.

Если принимается решение γ_0 или γ_1, то на этом эксперимент заканчивается. В противном случае извлекают следующую выборку, и описанная процедура повторяется: на основании выборки (x_1, x_2) размера $n=2$ либо принимается гипотеза H_0, либо она отвергается, либо считается, что эта выборка недостаточна, чтобы принять окончательное решение γ_0 или γ_1. Если принята или отвергнута гипотеза H_0, то на этом наблюдения оканчиваются, а если окончательное решение не принято, то наблюдают следующее значение x_3, и указанная процедура повторяется относительно выборки (x_1, x_2, x_3).

Испытание заканчивается на той выборке, на основании которой удается принять одно из двух решений, γ_0 или γ_1. При последовательном анализе размер выборки n заранее неизвестен и является *случайной величиной*. Теперь на каждом этапе пространство выборок соответствующего числа измерений должно делиться не на две, а на три области: критическую G_1, допустимую G_0 и промежуточную $G_{\text{пр}}$. Разделение пространства выборок на три области и содержит указание на то, должно ли быть принято одно из решений, γ_0 или γ_1, или эксперимент должен быть продолжен. Если выборочное значение попадает в критическую область G_1, то гипотеза H_0 отвергается; если — в допустимую область G_0, то она принимается, и если выборочное значение попало в промежуточную область $G_{\text{пр}}$, то это служит указанием на необходимость продолжить наблюдения.

Как и в случае непоследовательного анализа, число способов разбиения пространства выборок не ограничено. Следовательно, возможны самые разнообразные правила выбора решения, и, очевидно, опять необходимы какие-то критерии качества, при помощи которых можно было бы сравнивать различные процедуры последовательного анализа и выбирать наилучшую. В качестве такого критерия, вполне разумного и естественного, часто выбирают *минимальную среднюю стоимость эксперимента*. Если считать, что стоимость эксперимента пропорциональна размеру выборки n, то критерием качества последовательного правила выбора решения служит *минимум среднего значения* размера выборки, необходимой для принятия решения γ_0 или γ_1 (т. е. для завершения последовательного анализа) при условии, что уровень значимости не превышает α,

а мощность не меньше $1-\beta$. Заметим при этом, что средние значения размера выборки $m_1\{n \mid H_0\}$ и $m_1\{n \mid H_1\}$ при справедливости гипотез H_0 и H_1 соответственно, вообще говоря, не равны и требуется минимизация *обеих* величин.

Как показал А. Вальд [3], среди всех правил выбора решений (последовательных и непоследовательных), для которых условные вероятности ошибок не превосходят величин α и β, последовательное правило выбора решения, состоящее в сравнении *отношения правдоподобия* $l(x_1, \ldots, x_n)$ с двумя порогами *) c_0 и c_1, приводит к наименьшим значениям $m_1\{n \mid H_0\}$ и $m_1\{n \mid H_1\}$.

Оптимальное разбиение пространства выборок определяется следующими неравенствами:

для допустимой области (G_0)

$$c_0 < l(x_1, \ldots, x_k) < c_1, \quad k=1, \ldots, n-1, \qquad (1.55)$$
$$l(x_1, \ldots, x_n) \leqslant c_0;$$

для критической области (G_1)

$$c_0 < l(x_1, \ldots, x_k) < c_1, \quad k=1, \ldots, n-1, \qquad (1.55')$$
$$l(x_1, \ldots, x_n) \geqslant c_1;$$

для промежуточной области ($G_{\text{пр}}$)

$$c_0 < l(x_1, \ldots, x_k) < c_1, \quad k=1, \ldots, n. \qquad (1.55'')$$

Точное определение порогов c_0 и c_1 сопряжено со значительными математическими трудностями. Однако было доказано [3], что

$$c_1 \leqslant \max\left(\frac{\beta}{1-\alpha}, \frac{1-\beta}{\alpha}\right), \qquad (1.56)$$

$$c_0 \geqslant \min\left(\frac{\beta}{1-\alpha}, \frac{1-\beta}{\alpha}\right). \qquad (1.56')$$

В практически интересных случаях, когда условные вероятности ошибок не превышают значения 0,5, имеем $\frac{1-\beta}{\alpha} > \frac{\beta}{1-\alpha}$, и эти неравенства могут быть переписаны в виде

$$c_1 \leqslant \frac{1-\beta}{\alpha}, \quad c_0 \geqslant \frac{\beta}{1-\alpha}. \qquad (1.57)$$

*) Величины c_0 и c_1 зависят, вообще говоря, от n.

Следующий шаг состоит в замене неравенств (1.57) равенствами, что не очень сильно увеличивает погрешность (подробнее см. [7, стр. 142—143]).

Таким образом, с некоторым приближением оптимальное последовательное правило выбора решения может быть сформулировано следующим образом: при n-м наблюдении принимается решение γ_0, если *)

$$\frac{\beta}{1-\alpha} < l(x_1, \ldots, x_k) < \frac{1-\beta}{\alpha}, \quad k - 1, \ldots, n-1 \tag{1.58}$$

и

$$l(x_1, \ldots, x_n) \leqslant \frac{\beta}{1-\alpha} < 1; \tag{1.58'}$$

принимается решение γ_1, если выполняются неравенства (1.58) и

$$l(x_1, \ldots, x_n) \geqslant \frac{1-\beta}{\alpha} > 1. \tag{1.58''}$$

Заметим, что при указанных выше допущениях, в отличие от байесовского, последовательное правило предписывает сравнение отношения правдоподобия с порогами, не зависящими от априорных вероятностей состояний и от величин потерь. Эти пороги определяются заданными вероятностями ошибок первого и второго рода.

Если вместо отношения правдоподобия вычислять его логарифм, то указанное выше правило можно переписать в виде: при n-м наблюдении принимается решение γ_0, если

$$\ln\frac{\beta}{1-\alpha} < \ln l(x_1, \ldots, x_k) < \ln\frac{1-\beta}{\alpha}, \quad k - 1, \ldots, n-1 \tag{1.59}$$

и

$$\ln l(x_1, \ldots, x_n) \leqslant \ln\frac{\beta}{1-\alpha} < 0; \tag{1.59'}$$

принимается решение γ_1, если наряду с (1.59)

$$\ln l(x_1, \ldots, x_n) \geqslant \ln\frac{1-\beta}{\alpha} > 0. \tag{1.59''}$$

Определим теперь величины средних значений размера выборки $m_1\{n \mid H_0\}$ и $m_1\{n \mid H_1\}$, соответствующих после-

*) Вероятность того, что число этапов n в последовательной процедуре проверки гипотезы *не ограничено*, равна нулю (см. [7, стр. 166]).

довательному правилу выбора решения. В рассматриваемом случае логарифм отношения правдоподобия представляет сумму *случайного числа* случайных величин [см. (1.38)]:

$$\ln l(x_1, \ldots, x_n) = \sum_{i=1}^{n} \ln l(x_i).$$

Поэтому [см. (3.140) в первой книге]

$$m_1\{\ln l(x_1, \ldots, x_n) \mid H_{0;\,1}\} = m_1\{n \mid H_{0;\,1}\}\, m_1\{\ln l(x) \mid H_{0;\,1}\},$$

откуда

$$m_1\{n \mid H_0\} = \frac{m_1\{\ln l(x_1, \ldots, x_n) \mid H_0\}}{m_1\{\ln l(x) \mid H_0\}}, \qquad (1.60)$$

$$m_1\{n \mid H_1\} = \frac{m_1\{\ln l(x_1, \ldots, x_n) \mid H_1\}}{m_1\{\ln l(x) \mid H_1\}}. \qquad (1.60')$$

Предположим, что при принятии решения (γ_0 или γ_1) на n-м шаге величина отношения правдоподобия *точно совпадает* с одним из порогов, c_0 или c_1 (т. е. будем пренебрегать пересечением порога на заключительном этапе проверки гипотезы). Тогда $\ln l(x_1, \ldots, x_n)$ представляет дискретную случайную величину, принимающую два значения $\ln c_0$ и $\ln c_1$ с вероятностями $1-\alpha$ и α, если верна гипотеза H_0, и с вероятностями β и $1-\beta$, если верна гипотеза H_1. Отсюда следует:

$$m_1\{\ln l(x_1, \ldots, x_n) \mid H_0\} = (1-\alpha)\ln\frac{\beta}{1-\alpha} + \alpha \ln\frac{1-\beta}{\alpha}, \qquad (1.61)$$

$$m_1\{\ln l(x_1, \ldots, x_n) \mid H_1\} = \beta \ln\frac{\beta}{1-\alpha} + (1-\beta)\ln\frac{1-\beta}{\alpha}. \qquad (1.61')$$

Подставляя (1.61) в (1.60) и (1.61′) в (1.60′), получим

$$m_1\{n \mid H_0\} = \frac{(1-\alpha)\ln\dfrac{\beta}{1-\alpha} + \alpha\ln\dfrac{1-\beta}{\alpha}}{m_{10}}, \qquad (1.62)$$

$$m_1\{n \mid H_1\} = \frac{\beta\ln\dfrac{\beta}{1-\alpha} + (1-\beta)\ln\dfrac{1-\beta}{\alpha}}{m_{11}}, \qquad (1.62')$$

где [см. (1.47) и (1.48)]

$$m_{10} = m_1\{\ln l(x) \mid H_0\} = \int\limits_{-\infty}^{\infty} W_1(x \mid s_0) \ln \frac{W_1(x \mid s_1)}{W_1(x \mid s_0)} dx; \quad (1.63)$$

$$m_{11} = m_1\{\ln l(x) \mid H_1\} = \int\limits_{-\infty}^{\infty} W_1(x \mid s_1) \ln \frac{W_1(x \mid s_1)}{W_1(x \mid s_0)} dx, \quad (1.63')$$

причем $m_{10} < 0$, $m_{11} > 0$, что непосредственно следует из формулы (23) задачи 1.10.

Средние значения объема выборки, определяемые по формулам (1.62) и (1.62′), являются *минимально возможными*,

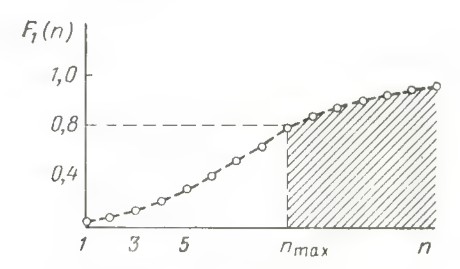


Рис. 3. Функция распределения размера выборки при последовательном анализе.

если рассматривать любые другие правила выбора решений (*в том числе и непоследовательные*), гарантирующие ограничение вероятностей ошибок заданными величинами. В этом и состоит основное достоинство последовательного анализа, так как его использование позволяет в среднем сэкономить на продолжительности и стоимости эксперимента. Однако оговорка, что эта *экономия достигается в среднем*, имеет весьма существенное значение. Так как в последовательном анализе размер выборки n — величина случайная и ее возможные значения могут быть значительно больше среднего значения, то в данном конкретном эксперименте может оказаться, что оптимальная последовательная процедура приведет к чрезмерно большому размеру выборки.

Это обстоятельство иллюстрирует рис. 3. Заштрихованная часть соответствует размерам выборок, большим неко-

торой допустимой величины n_{max}. Хотя в 80% случаев будет иметь место выигрыш в длительности процедуры проверки гипотезы по сравнению с непоследовательным анализом с заданным размером $n = n_{max}$, все-таки будут такие «несчастливые» случаи, когда процедура последовательного анализа окажется более длительной, чем непоследовательного. Естественно, приходит мысль о способе устранения этого недостатка, который заключается в том, что заранее устанавливается максимальное значение объема выборки n_{max}, при достижении которого процедура последовательного анализа заканчивается и обязательно принимается одно из решений, γ_0 или γ_1 (если, конечно, они не были приняты раньше). Таким образом, можно обезопасить себя от случаев, когда $n > n_{max}$.

Указанная процедура проверки гипотез называется *усеченным* последовательным анализом. При усеченном последовательном анализе до n_{max} устанавливается два порога, с которыми сравнивается отношение правдоподобия, а если размер выборки достигает величины n_{max}, то соответствующее отношение правдоподобия сравнивается не с двумя порогами, а только с одним, как при непоследовательном анализе. Чем меньше n_{max}, т. е. чем сильнее усечение, тем меньшим будет выигрыш в среднем времени, получаемом от последовательного анализа.

Правило выбора решения при усеченном последовательном анализе формулируется следующим образом: если при размерах выборки $n \leqslant n_{max}$ правило (1.59) не приводит к выбору одного из решений (γ_0 или γ_1), то гипотеза H_0 отклоняется (принимается решение γ_1), если

$$\ln\frac{\beta}{1-\alpha} < \ln c \leqslant \ln l\,(x_1, \ldots, x_{n_{max}}) < \ln\frac{1-\beta}{\alpha}, \qquad (1.64)$$

и принимается гипотеза H_0 (решение γ_0), если

$$\ln\frac{\beta}{1-\alpha} < \ln l\,(x_1, \ldots, x_{n_{max}}) < \ln c < \ln\frac{1-\beta}{\alpha}. \qquad (1.64')$$

При использовании правила (1.64) вероятности ошибок $\alpha_{ус}$, $\beta_{ус}$ первого и второго рода могут оказаться большими, чем заданные $\alpha_{ус} \geqslant \alpha$, $\beta_{ус} \geqslant \beta$, так как принимаемые при этом ошибочные решения, возможно, не появились бы при продолжении испытаний ($n > n_{max}$). Верхние границы условных вероятностей ошибок при усеченной последовательной процедуре принятия решения определяются

неравенствами

$$\alpha_{ус} \leqslant \alpha + P\left\{\ln\frac{1-\beta}{\alpha} > \ln l(x_1, \ldots, x_{n_{max}}) \geqslant \ln c \mid H_0\right\}, \tag{1.65}$$

$$\beta_{ус} \leqslant \beta + P\left\{\ln c > \ln l(x_1, \ldots, x_{n_{max}}) > \ln\frac{\beta}{1-\alpha} \,\Big|\, H_1\right\}. \tag{1.65'}$$

При $n_{max} \gg 1$, как показано в [3], эти неравенства записываются в виде

$$\alpha_{ус} \leqslant \alpha + F\left[\frac{\ln\frac{1-\beta}{\alpha} - n_{max}m_{10}}{\sqrt{n_{max}M_{20}}}\right] - F\left[\frac{\ln c - n_{max}m_{10}}{\sqrt{n_{max}M_{20}}}\right],$$

$$\beta_{ус} \leqslant \beta + F\left[\frac{\ln c - n_{max}m_{11}}{\sqrt{n_{max}M_{21}}}\right] - F\left[\frac{\ln\frac{\beta}{1-\alpha} - n_{max}m_{11}}{\sqrt{n_{max}M_{21}}}\right],$$

где m_{10}, m_{11}, M_{20}, M_{21} определяются согласно (1.47) и (1.48) и $F(x)$ — интеграл Лапласа.

Заметим, что так как $\ln\frac{\beta}{1-\alpha} < 0$ и $\ln\frac{1-\beta}{\alpha} > 0$, то величина порога $c = 1$ при усеченной последовательной процедуре всегда приемлема [см. (1.64), (1.64')]. Соображений относительно оптимального выбора порога c пока не существует.

1.3. ПРОВЕРКА ПРОСТОЙ ГИПОТЕЗЫ О ПАРАМЕТРЕ РАСПРЕДЕЛЕНИЯ

1.3.1. Среднее значение нормальной случайной величины. Проиллюстрируем теперь изложенные в этом разделе общие методы на нескольких примерах. Начнем с простейшего примера проверки гипотезы H_0 о том, что среднее значение нормальной случайной величины равно a_0, при простой альтернативе H_1, что среднее равно a_1. Предположим, что дисперсия σ^2 нормальной случайной величины известна точно и что выборка, на основании которой проверяется гипотеза, состоит из независимых элементов $x_1, \ldots, x_n$.

Так как любая из рассмотренных оптимальных процедур проверки гипотезы состоит в сравнении отношения правдоподобия $l(x_1, \ldots, x_n)$ или его логарифма с некоторым порогом (или с двумя порогами при последовательном анализе), то выпишем сначала выражение для $\ln l(x_1, \ldots, x_n)$ в случае нормального распределения при независимых элементах выборки:

$$\ln l(x_1, \ldots, x_n) = \sum_{i=1}^{n} \ln l(x_i)$$

$$= \sum_{i=1}^{n} \ln \frac{\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - a_1)^2}{2\sigma^2}\right]}{\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - a_0)^2}{2\sigma^2}\right]} \quad \frac{1}{2\sigma^2} \sum_{i=1}^{n} [(x_i -$$

$$- a_0)^2 - (x_i - a_1)^2] \quad \frac{a_1 - a_0}{\sigma^2} \sum_{i=1}^{n} x_i - \frac{n(a_1^2 - a_0^2)}{2\sigma^2}. \quad (1.66)$$

Теперь нетрудно сформулировать правило выбора решения (при непоследовательном анализе, т. е. фиксированном заранее размере n выборки). Принимается решение γ_1 (среднее равно a_1), если для наблюдаемой выборки $x_1, \ldots, x_n$

$$\frac{1}{n} \sum_{i=1}^{n} x_i \geqslant \frac{a_0 + a_1}{2} + \frac{\sigma^2 \ln c}{n(a_1 - a_0)} = K, \; a_1 > a_0, \quad (1.67)$$

и принимается решение γ_0 (среднее равно a_0), если выполняется неравенство, противоположное (1.67). Величина порога в *зависимости от выбранного критерия* определяется согласно табл. 1 на стр. 40.

Если $a_1 < a_0$, то знак неравенства (1.67) изменяется на противоположный и правило выбора решения γ_1 записывается в виде

$$\frac{1}{n} \sum_{i=1}^{n} x_i \leqslant \frac{a_0 + a_1}{2} - \frac{\sigma^2 \ln c}{n(a_0 - a_1)}, \; a_1 < a_0. \quad (1.67')$$

Таким образом, процедура проверки гипотезы о среднем значении нормальной случайной величины сводится к сравнению среднего арифметического выборочных значений с порогом:

$$K = \frac{a_0 + a_1}{2} + \frac{\sigma^2 \ln c}{n(a_1 - a_0)}. \quad (1.68)$$

Поверхность, разделяющая допустимую и критическую области пространства выборок, представляет в этом случае гиперплоскость, перпендикулярную единичному вектору и расположенную от начала координат на расстоянии, равном nK.

Нетрудно в рассматриваемом случае найти выражения условных вероятностей ошибок. Так как сумма нормальных случайных величин имеет нормальное распределение и так как

$$m_1\left\{\frac{1}{n}\sum_{i=1}^{n} x_i \mid H_0\right\} = a_0, \quad m_1\left\{\frac{1}{n}\sum_{i=1}^{n} x_i \mid H_1\right\} = a_1, \tag{1.69}$$

а при независимых элементах выборки

$$M_2\left\{\frac{1}{n}\sum_{i=1}^{n} x_i \mid H_0\right\} = M_2\left\{\frac{1}{n}\sum_{i=1}^{n} x_i \mid H_1\right\} = \frac{\sigma^2}{n}, \tag{1.70}$$

то, используя (1.67) [см. также (1.43) и (1.44)], находим (при $a_1 > a_0$)

$$\alpha = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\int_K^{\infty} e^{-\frac{n(y-a_0)^2}{2\sigma^2}}\,dy =$$

$$= \frac{1}{\sqrt{2\pi}}\int_{\frac{K-a_0}{\sigma}\sqrt{n}}^{\infty} e^{-\frac{t^2}{2}}\,dt = 1 - F\left(\frac{K-a_0}{\sigma}\sqrt{n}\right), \tag{1.71}$$

$$\beta = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\int_{-\infty}^{K} e^{-\frac{n(y-a_1)^2}{2\sigma^2}}\,dy =$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{K-a_1}{\sigma}\sqrt{n}} e^{-\frac{t^2}{2}}\,dt = F\left(\frac{K-a_1}{\sigma}\sqrt{n}\right), \tag{1.72}$$

где $F(x)$ — интеграл Лапласа [см. (1.49) и (1.50)].

На рис. 4 иллюстрируется рассматриваемое правило выбора решения. Пунктирные кривые изображают исходные функции распределения, соответствующие гипотезам H_0 и H_1. Сплошными линиями показаны функции распределения среднего арифметического выборочных значений для тех же гипотез. Порог установлен по критерию максимального правдоподобия ($c = 1$). Площади заштрихованных

участков равны вероятностям ошибок первого и второго рода, которые при $K = \frac{a_0 + a_1}{2}$ совпадают.

Величина K в (1.71) и (1.72) для первых трех критериев, указанных в табл. 1 (на стр. 40), определяется непосредст-

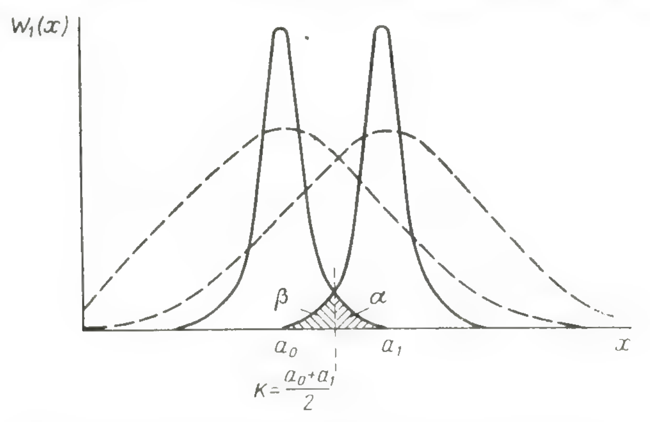


Рис. 4. Правило выбора решения для проверки гипотезы о среднем нормальной случайной величины.

венно по формуле (1.68), так что эти формулы могут быть переписаны в виде

$$\alpha = 1 - F\left[\frac{a_1 - a_0}{2\sigma}\sqrt{n} + \frac{\sigma \ln c}{\sqrt{n}\,(a_1 - a_0)}\right], \qquad (1.73)$$

$$\beta = F\left[-\frac{a_1 - a_0}{2\sigma}\sqrt{n} + \frac{\sigma \ln c}{\sqrt{n}\,(a_1 - a_0)}\right], \quad a_1 > a_0. \qquad (1.73')$$

Из приведенных формул видно, что условные вероятности ошибок при заданном c определяются еще только *одной* величиной $d_n = \frac{a_1 - a_0}{\sigma}\sqrt{n}$. Нетрудно проверить, что среднее и дисперсия логарифма отношения правдоподобия связаны с этой величиной простыми соотношениями:

$$m_1\{\ln l(x_1, \ldots, x_n) \mid H_1\} =$$
$$= -m_1\{\ln l(x_1, \ldots, x_n) \mid H_0\} = \frac{d_n^2}{2},$$

$$M_2\{\ln l(x_1, \ldots, x_n) \mid H_1\} = M_2\{\ln l(x_1, \ldots, x_n) \mid H_0\} = d_n^2.$$

Заметим, что из (1.49) и (1.50) следует, что при $n \to \infty$ вероятности ошибок стремятся к нулю в соответствии с указан-

ным выше (§ 1.2.5) результатом. В этом случае распределение среднего арифметического выборочных значений приближается к дельта-функциям в точках $y = a_0$, $y = a_1$, а порог $K \to \frac{a_0 + a_1}{2}$.

Для критерия *максимального правдоподобия* (рис. 4)

$$\alpha = \beta = 1 - F\left(\frac{a_1 - a_0}{2\sigma}\sqrt{n}\right). \tag{1.74}$$

При заданном $\alpha = \beta$ из (1.74) находим необходимый размер выборки:

$$n = 4\left(\frac{\sigma}{a_1 - a_0}\right)^2 x_\alpha^2, \tag{1.74'}$$

где

$$x_\alpha = \arg F(x).$$

В математической статистике величину x_α называют обычно процентным отклонением случайной величины *), т. е. такую абсциссу кривой распределения, которая характеризуется тем, что часть площади под этой кривой, находящаяся правее точки x_α, равна α. Иначе говоря,

$$P\{\xi \geqslant x_\alpha\} = \alpha. \tag{1.75}$$

В формуле (1.74') величина x_α представляет, таким образом, процентное отклонение *нормальной* случайной величины. В приложении 1 дана таблица процентных точек нормального распределения.

Для критерия Неймана — Пирсона при заданном уровне значимости α величина K определяется по формуле ($a_1 > a_0$)

$$K = a_0 + \frac{\sigma}{\sqrt{n}} x_\alpha. \tag{1.76}$$

---

*) Очевидно, что процентное отклонение случайной величины ξ совпадает с $(1 - \alpha)$-*квантилью* ее функции распределения $F_1(x)$, так как согласно (1.75) x_α является корнем уравнения $F_1(x) = 1 - \alpha$ (см. первую книгу, стр. 48). Особые случаи могут возникать, когда $F(x)$ разрывна или когда имеются интервалы на действительной оси, для которых интегральная функция распределения сохраняет постоянное значение. Во избежание недоразумений следует иметь в виду, что иногда (см., например [5]) отклонение определяется иначе: в (1.75) вместо ξ подставляют ее абсолютное значение $P\{|\xi| \geqslant \lambda_\alpha\} = \alpha$. В этой книге везде используется определение (1.75). Для случайной величины с симметричной относительно начала координат функцией распределения $x_\alpha = \lambda_{2\alpha}$.

Вероятность ошибки второго рода в рассматриваемом случае определяется подстановкой (1.76) в (1.72):

$$\beta = F\left[-\frac{a_1 - a_0}{\sigma}\sqrt{n} + x_\alpha\right] \qquad (1.77)$$

или

$$x_\alpha - x_{1-\beta} = \frac{a_1 - a_0}{\sigma}\sqrt{n}, \quad a_1 > a_0. \qquad (1.77')$$

Соотношение (1.77′) устанавливает связь между заданным размером выборки n, условной вероятностью ошибки первого рода α и минимально возможной вероятностью ошибки второго рода β. При $n \to \infty$ из (1.77) следует, что $\beta \to 0$ в соответствии с общим результатом, указанным в § 1.2.5.

Заметим, что порог K, устанавливаемый по критерию Неймана — Пирсона в соответствии с (1.76), *не зависит* от величины a_1.

Следует подчеркнуть, что вероятности ошибок α и β в байесовском решении и вероятность ошибки второго рода (при заданном α) для критерия Неймана — Пирсона зависят не от каждой из величин n, a_0, a_1 и σ в отдельности, а лишь от их единственной комбинации $d_n = \frac{a_1 - a_0}{\sigma}\sqrt{n}$. Отсюда следует, что при уменьшении (в K раз) величины $\frac{a_1 - a_0}{\sigma}$ (различение близких гипотез) для сохранения величин вероятности ошибок потребуется увеличение (в K^2 раз) размера выборки n. Состоятельность правила решения будет в этом случае лишь тогда, когда $\frac{a_1 - a_0}{\sigma}$ убывает медленнее, чем $\frac{1}{\sqrt{n}}$ при $n \to \infty$.

Если $a_1 < a_0$, то решение γ_1 по критерию Неймана — Пирсона принимается при условии [ср. (1.67′)]

$$\frac{1}{n}\sum_{i=1}^{n} x_i \leqslant a_0 - \frac{\sigma}{\sqrt{n}} x_\alpha.$$

Вероятность ошибки второго рода тогда равна

$$\beta = F\left[-\frac{a_0 - a_1}{\sigma}\sqrt{n} + x_\alpha\right].$$

Рассмотрим теперь *минимаксное правило* решения, полагая $\Pi_{00} = \Pi_{11} = 0$, $\Pi_{01} = \lambda\Pi_{10}$. Тогда $c^* = \lambda$, $c = \lambda\mu = \frac{q\lambda}{p}$

и уравнение, определяющее наименее благоприятную величину априорной вероятности $q_{\text{мм}} = 1 - p_{\text{мм}}$, может быть в соответствии с (1.51) и (1.73) и (1.73′) представлено в виде

$$1 - F\left[\frac{a_1 - a_0}{2\sigma}\sqrt{n} + \frac{\sigma}{\sqrt{n}\,(a_1 - a_0)}\ln\frac{q\lambda}{1-q}\right] =$$

$$= \frac{1}{\lambda} F\left[-\frac{a_1 - a_0}{2\sigma}\sqrt{n} + \frac{\sigma}{\sqrt{n}\,(a_1 - a_0)}\ln\frac{q\lambda}{1-q}\right]. \quad (1.78)$$

При $\lambda = 1$ из (1.78) следует, что $q_{\text{мм}} = \frac{1}{2}$. Как и следовало ожидать, при одинаковых потерях из-за ошибок первого

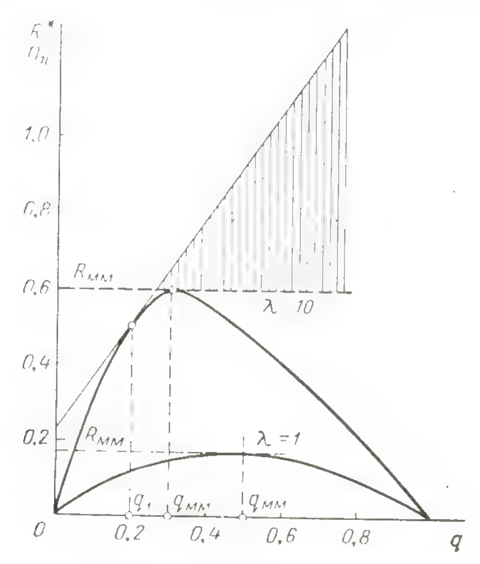


Рис. 5. Средний риск.

и второго рода наименее благоприятным с точки зрения байесовского риска является равновероятное появление состояний s_0 и s_1.

На рис. 5 показана зависимость байесовского риска R^* от величины q при $\frac{a_1 - a_0}{2\sigma}\sqrt{n} = 1$. Для этого случая, как это следует из (1.23), (1.73) и (1.73′), выражение для R^* имеет вид

$$R^* = \Pi_{10}\left\{\lambda q\left[1 - F\left(\ln\frac{q\lambda}{1-q} + 1\right)\right] + pF\left(\ln\frac{q\lambda}{1-q} - 1\right)\right\}.$$

При заданном λ наименее благоприятная величина $q_{мм}$ соответствует максимальному значению $R_{мм} = R^*(q_{мм})$. Как видно из рис. 5, при $\lambda = 1$ потери уменьшаются не очень значительно при $q \neq q_{мм}$, однако при $\lambda = 10$ минимаксное правило может показаться чересчур осторожным. Такая осторожность, однако, является гарантией того, что потери никогда не превысят величины $R_{мм}$. Действительно, если при $\lambda = 10$ немного отойти от наименее благоприятной величины $q_{мм} = 0,3$ и принять байесовское решение при $q_1 = 0,2$, то средние потери будут уменьшены всего на 20%. Если же в действительности $q \neq q_1$, а применяется *байесовское* решение для $q = q_1$, то средний риск будет изменяться в зависимости от q по линейному закону $R(q) = qr_0(q_1) + (1 - q)\,r_1(q_1)$ (касательная в точке $q = q_1$ к кривой $R^*(q)$ при $\lambda = 10$ на рис. 5) и при некоторых значениях q может *значительно* превысить величину $R_{мм}$, соответствующую минимаксному правилу (см. заштрихованную часть рис. 5).

В заключение этого примера рассмотрим *последовательное* правило выбора решения. В этом случае для заданных вероятностей ошибок α и β логарифм отношения правдоподобия (1.66) должен сравниваться с двумя порогами $\ln \frac{\beta}{1-\alpha}$ и $\ln \frac{1-\beta}{\alpha}$ (при $\alpha < 0,5$, $\beta < 0,5$). Подробнее говоря, при $a_1 > a_0$ на n-м этапе принимается гипотеза, что среднее нормальной случайной величины равно a_0, если

$$\frac{1}{n}\sum_{i=1}^{n} x_i \leqslant \frac{a_0 + a_1}{2} + \frac{\sigma^2 \ln \frac{\beta}{1-\alpha}}{n\,(a_1 - a_0)}, \qquad (1.79)$$

принимается гипотеза, что среднее равно a_1, если

$$\frac{1}{n}\sum_{i=1}^{n} x_i \geqslant \frac{a_0 + a_1}{2} + \frac{\sigma^2 \ln \frac{1-\beta}{\alpha}}{n\,(a_1 - a_0)}, \qquad (1.79')$$

и продолжаются наблюдения, если не выполняется ни одно из неравенств (1.79) и (1.79′).

Для того чтобы определить средние значения размеров выборки до принятия окончательного решения (минимально возможные), необходимо в соответствии с (1.62) и (1.62′) определить средние логарифма одномерного отношения правдоподобия для двух случаев: 1) когда параметр рас-

пределения равен a_0, 2) когда этот параметр равен a_1. Так как согласно (1.66)

$$\ln l(x) = \frac{a_1 - a_0}{\sigma^2} x - \frac{a_1^2 - a_0^2}{2\sigma^2},$$

где x — нормальная случайная величина с известной дисперсией σ^2 и средним a_0 (или a_1), то в рассматриваемом примере

$$m_{10} = \frac{a_1 - a_0}{\sigma^2} a_0 - \frac{a_1^2 - a_0^2}{2\sigma^2} = -\frac{(a_1 - a_0)^2}{2\sigma^2}, \qquad (1.80)$$

$$m_{11} = \frac{a_1 - a_0}{\sigma^2} a_1 - \frac{a_1^2 - a_0^2}{2\sigma^2} = \frac{(a_1 - a_0)^2}{2\sigma^2}. \qquad (1.81)$$

Подставляя (1.80) и (1.81) в (1.62) и (1.62′), получим

$$m_1\{n \mid a_0\} = 2\sigma^2 \frac{(1-\alpha)\ln\frac{1-\alpha}{\beta} + \alpha \ln\frac{\alpha}{1-\beta}}{(a_1 - a_0)^2}, \qquad (1.82)$$

$$m_1\{n \mid a_1\} = 2\sigma^2 \frac{\beta \ln\frac{\beta}{1-\alpha} + (1-\beta)\ln\frac{1-\beta}{\alpha}}{(a_1 - a_0)^2}. \qquad (1.83)$$

Из этих формул видно, что средний размер выборок до принятия окончательного решения (γ_0 или γ_1) обратно пропорционален квадрату величины $\frac{a_1 - a_0}{\sigma}$.

При $\alpha = \beta = \varepsilon < 0{,}5$

$$m_1\{n \mid a_0\} = m_1\{n \mid a_1\} = 2\left(\frac{\sigma}{a_1 - a_0}\right)^2 (1 - 2\varepsilon)\ln\frac{1-\varepsilon}{\varepsilon}. \qquad (1.84)$$

В табл. 2 приведены $\left(\text{при } \frac{a_1 - a_0}{\sigma} = 1\right)$ значения среднего объема выборки, подсчитанного согласно (1.84), и объема

Таблица 2

ε	0,001	0,01	0,05	0,1	0,2
$m_1\{n\}$	13,7	9,0	5,3	3,5	1,67
n	38,2	21,6	10,6	6,5	2,8
$n/m_1\{n\}$	2,78	2,41	2,00	1,85	1,68
$(\varepsilon_{ус})_{max}$	0,0023	0,0197	0,096	0,185	0,355

выборки, необходимого для достижения заданных условных вероятностей $\alpha = \beta = \varepsilon$ при использовании непоследовательного правила по критерию максимального правдоподобия [см. (1.74′)]. Предпоследняя строка этой таблицы указывает на величину выигрыша при использовании последовательного правила *)

$$\frac{n}{m_1\{n\}} = \frac{2x_\varepsilon^2}{(1-2\varepsilon)\ln\frac{1-\varepsilon}{\varepsilon}}, \qquad (1.84')$$

причем это отношение не зависит от величины $\frac{\sigma}{a_1-a_0}$.

Заметим, что для этого случая при усечении последовательного анализа $n \leqslant n_{max}$ и $c = 1$ из (1.65) и (1.65′) следует, что верхние границы вероятностей ошибок одинаковы и равны

$$(\varepsilon_{yc})_{max} = \varepsilon + F\left(\frac{d}{2} + \frac{1}{d}\ln\frac{1-\varepsilon}{\varepsilon}\right) - F\left(\frac{d}{2}\right), \qquad (1.84'')$$

где

$$d^2 = n_{max}\left(\frac{a_1-a_0}{\sigma}\right)^2.$$

В последней строке табл. 2 приведены величины $(\varepsilon_{yc})_{max}$ при $\frac{a_1-a_0}{\sigma} = 1$ и n_{max}, совпадающем с размером n выборки для непоследовательного правила (третья строка таблицы).

1.3.2. Дисперсия нормальной случайной величины. Предположим теперь, что среднее значение нормальной случайной величины равно нулю, а относительно ее дисперсии выдвигается гипотеза H_0, что она равна σ_0^2, против простой альтернативы H_1, что дисперсия равна σ_1^2. Сохраним предположение о независимости элементов выборки, на основании которой должно быть принято решение о принятии или отклонении гипотезы, и запишем выражение логарифма отношения правдоподобия:

$$\ln l(x_1, \ldots, x_n) = \sum_{k=1}^{n} \ln\left[\frac{\sigma_0}{\sigma_1}\exp\left\{-\frac{x_k^2}{2}\left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_0^2}\right)\right\}\right] =$$
$$= n\ln\frac{\sigma_0}{\sigma_1} - \frac{1}{2}\left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_0^2}\right)\sum_{k=1}^{n} x_k^2. \qquad (1.85)$$

*) Здесь x_ε — процентное отклонение нормальной случайной величины. Заметим также, что рассмотренный для примера случай $\alpha = \beta$ не соответствует наибольшей эффективности последовательного анализа. Величины выигрыша получаются большими при $\alpha \ll \beta$ или $\beta \ll \alpha$ (см. [1, стр. 57]).

Тогда правило выбора решения (при непоследовательном анализе) можно записать следующим образом: принимается решение γ_1 (дисперсия равна σ_1^2), если при $\sigma_1 > \sigma_0$

$$\sum_{k=1}^{n} x_k^2 \geqslant \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2-\sigma_0^2} \ln\left[\left(\frac{\sigma_1}{\sigma_0}\right)^n c\right] = K, \tag{1.86}$$

и принимается решение γ_0 (дисперсия равна σ_0^2), если выполняется неравенство, противоположное (1.86).

Таким образом, процедура проверки гипотезы о дисперсии нормальной случайной величины сводится к сравнению *суммы квадратов* выборочных значений с порогом

$$K = \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2-\sigma_0^2} \ln\left[\left(\frac{\sigma_1}{\sigma_0}\right)^n c\right] > 0, \quad \sigma_1 > \sigma_0. \tag{1.86'}$$

Поверхность, разделяющая допустимую и критическую области пространства выборок, представляет в этом случае гиперсферу с центром в начале координат и радиусом, равным $\sqrt{K}$.

Используя (1.25) и (1.26), можно найти выражения условных вероятностей ошибок. Так как сумма $\sum_{k=1}^{n}\left(\frac{x_k}{\sigma}\right)^2$ квадратов нормированных нормальных случайных величин имеет χ^2-распределения с n степенями свободы (см. первую книгу, стр. 164), то плотности вероятности суммы квадратов в левой части (1.86) равны

$$w_{10}(y) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)\sigma_0^n}\left(\frac{y}{2}\right)^{\frac{n}{2}-1} e^{-\frac{y}{2\sigma_0^2}}, \quad y > 0, \tag{1.87}$$

если верна гипотеза H_0, и

$$w_{11}(y) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)\sigma_1^n}\left(\frac{y}{2}\right)^{\frac{n}{2}-1} e^{-\frac{y}{2\sigma_1^2}}, \quad y > 0, \tag{1.87'}$$

если верна гипотеза H_1. Тогда

$$\alpha = P\left\{\sum_{k=1}^{n} x_k^2 \geqslant K \mid H_0\right\} = \int_K^{\infty} w_{10}(y)\,dy =$$

$$= 1 - \frac{\Gamma\left(\frac{n}{2}, \frac{K}{2\sigma_0^2}\right)}{\Gamma\left(\frac{n}{2}\right)}, \tag{1.88}$$

$$\beta = P\left\{\sum_{k=1}^{n} x_k^2 < K \mid H_1\right\} = \int_0^K w_{11}(y)\,dy =$$

$$= \frac{\Gamma\left(\frac{n}{2}, \frac{K}{2\sigma_1^2}\right)}{\Gamma\left(\frac{n}{2}\right)}, \qquad (1.88')$$

где $\Gamma(u, v)$ — неполная гамма-функция.

Рис. 6 иллюстрирует правило выбора решения (1.86). Пунктирные кривые изображают исходные функции рас-

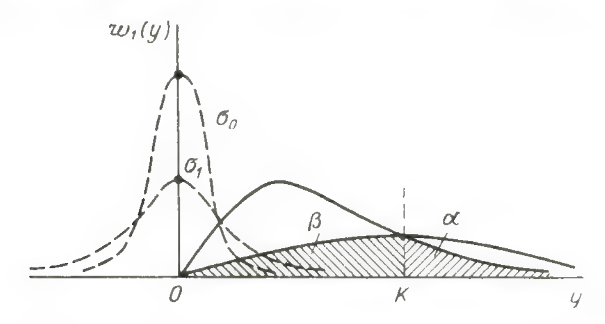


Рис. 6. Правило выбора решения для проверки гипотезы о дисперсии нормальной случайной величины.

пределения, соответствующие гипотезам H_0 и H_1. Сплошными линиями показаны распределения суммы квадратов выборочных значений для тех же гипотез. Площади заштрихованных участков равны вероятностям ошибок первого и второго рода.

Величина K в (1.88) и (1.88′) для первых трех критериев, указанных в табл. 1 (на стр. 40), определяется непосредственно из (1.86′). Для критерия Неймана — Пирсона при заданном α из (1.88) следует:

$$K = \sigma_0^2 \chi_\alpha^2, \qquad (1.89)$$

где χ_α^2 — процентное отклонение случайной величины, распределенной по закону χ^2 с n степенями свободы *). Заметим, что порог K, устанавливаемый согласно (1.89), *не зависит* от величины σ_1.

*) Таблица процентных точек χ^2-распределения приведена в приложении II.

Подставляя (1.89) в (1.88′), получим соотношение

$$\frac{\chi_\alpha^2}{\chi_{1-\beta}^2}=\frac{\sigma_1^2}{\sigma_0^2}, \qquad (1.89')$$

связывающее заданную вероятность ошибки первого рода α, минимально возможную вероятность ошибки второго рода β и размер выборки n, который неявно входит в (1.89′) через χ_α^2 и $\chi_{1-\beta}^2$. Отсюда следует, что необходимый размер выборки для достижения величины β (при заданном α) зависит от отношения дисперсий σ_1^2/σ_0^2.

Если $n \gg 1$, можно использовать асимптотическую нормальность случайной величины, имеющей χ^2-распределение [(см. первую книгу, формула (1.43)], и выразить величины α и β через интеграл Лапласа [см. также (1.49) и (1.50)]:

$$\alpha \sim 1-F\left(\frac{\sqrt{K}}{\sigma_0}-\sqrt{2n}\right), \qquad (1.90)$$

$$\beta \sim F\left(\frac{\sqrt{K}}{\sigma_1}-\sqrt{2n}\right). \qquad (1.90')$$

При $n\to\infty$ в соответствии с общим результатом, указанным в § 1.2.5, вероятности ошибок стремятся к нулю, так как при $\sigma_1>\sigma_0$ имеем

$$\frac{\sqrt{K}}{\sigma_0}-\sqrt{2n}\sim\sqrt{2n}\left[\frac{\frac{\sigma_1}{\sigma_0}\sqrt{\ln\frac{\sigma_1}{\sigma_0}}}{\sqrt{\left(\frac{\sigma_1}{\sigma_0}\right)^2-1}}-1\right]\to\infty,$$

$$\frac{\sqrt{K}}{\sigma_1}-\sqrt{2n}\sim\sqrt{2n}\left[\frac{\sqrt{\ln\frac{\sigma_1}{\sigma_0}}}{\sqrt{\left(\frac{\sigma_1}{\sigma_0}\right)^2-1}}-1\right]\to-\infty.$$

Заметим, что при различении близких гипотез $\left(\frac{\sigma_1}{\sigma_0}\to 1\right)$ размер выборки растет пропорционально $\frac{\sigma_0^2}{\sigma_1^2-\sigma_0^2}$.

Для минимаксного правила решения при $\Pi_{00}=\Pi_{11}=0$, $\Pi_{01}=\lambda\Pi_{10}$ уравнение, определяющее наименее благоприятную величину априорной вероятности $q_{\text{мм}}$, может быть в соответствии с (1.51), (1.86′), (1.88) и (1.88′) представ-

лено в виде ($\sigma_1 > \sigma_0$)

$$\Gamma\left(\frac{n}{2}\right) - \Gamma\left\{\frac{n}{2}, \frac{\sigma_1^2}{\sigma_1^2-\sigma_0^2}\ln\left[\left(\frac{\sigma_1}{\sigma_0}\right)^n \frac{q\lambda}{p}\right]\right\} =$$

$$= \frac{1}{\lambda}\Gamma\left\{\frac{n}{2}, \frac{\sigma_0^2}{\sigma_1^2-\sigma_0^2}\ln\left[\left(\frac{\sigma_1}{\sigma_0}\right)^n \frac{q\lambda}{p}\right]\right\}, \qquad (1.91)$$

а при $n \gg 1$ (выражая неполную гамма-функцию через интеграл Лапласа) — в виде

$$\frac{1-F\,(c_1\sqrt{2n})}{1-F\,(c_2\sqrt{2n})} = \lambda, \qquad (1.91')$$

где

$$c_1 = 1 - \frac{1}{\sqrt{\left(\frac{\sigma_1}{\sigma_0}\right)^2 - 1}}\left[\ln\frac{\sigma_1}{\sigma_0} + \frac{1}{n}\ln\frac{q\lambda}{p}\right]^{1/2};$$

$$c_2 = \frac{\frac{\sigma_1}{\sigma_0}}{\sqrt{\left(\frac{\sigma_1}{\sigma_0}\right)^2 - 1}}\left[\ln\frac{\sigma_1}{\sigma_0} + \frac{1}{n}\ln\frac{q\lambda}{p}\right]^{1/2} - 1. \qquad (1.91'')$$

Из (1.91′) следует, что (в отличие от минимаксного правила при проверке гипотезы о среднем) при $\lambda = 1$ наименее благоприятная величина $q_{\text{мм}} \neq 1/2$. Эта величина находится (при $\lambda = 1$) из равенства $c_1 = c_2$, которое с учетом (1.91″) приводится к виду

$$\mu_{\text{мм}} = \frac{q_{\text{мм}}}{p_{\text{мм}}} = \left(\frac{\sigma_0}{\sigma_1}\right)^n \exp\left[4n\frac{\sigma_1-\sigma_0}{\sigma_1+\sigma_0}\right].$$

Переходим к последовательному правилу выбора решений. Для заданных вероятностей ошибок α и β оно в соответствии с (1.59) формулируется следующим образом: при $\sigma_1 > \sigma_0$ принимается гипотеза, что дисперсия нормальной случайной величины равна σ_0^2, если для наблюдаемой выборки

$$\sum_{k=1}^{n} x_k^2 \leqslant \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2-\sigma_0^2}\ln\left[\left(\frac{\sigma_1}{\sigma_0}\right)^n \frac{\beta}{1-\alpha}\right], \qquad (1.92)$$

принимается гипотеза, что дисперсия равна σ_1^2, если

$$\sum_{k=1}^{n} x_k^2 \geqslant \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2-\sigma_0^2}\ln\left[\left(\frac{\sigma_1}{\sigma_0}\right)^n \frac{1-\beta}{\alpha}\right], \qquad (1.92')$$

и продолжаются наблюдения, если не выполняется ни одно из неравенств (1.92) и (1.92′).

Определим среднее значение размера выборки, необходимого для принятия окончательного решения, используя (1.62) и (1.62′). Из (1.85) имеем

$$m_{10} = m_1\{\ln l(x) \mid \sigma_0^2\} = -\frac{\sigma_1^2-\sigma_0^2}{2\sigma_0^2\sigma_1^2} m_1\{x^2 \mid \sigma_0^2\} + \ln\frac{\sigma_0}{\sigma_1} = \frac{\sigma_1^2-\sigma_0^2}{2\sigma_1^2} + \ln\frac{\sigma_0}{\sigma_1}, \quad (1.93)$$

$$m_{11} = m_1\{\ln l(x) \mid \sigma_1^2\} = \frac{\sigma_1^2-\sigma_0^2}{2\sigma_0^2\sigma_1^2} m_1\{x^2 \mid \sigma_1^2\} + \ln\frac{\sigma_0}{\sigma_1} = \frac{\sigma_1^2-\sigma_0^2}{2\sigma_0^2} + \ln\frac{\sigma_0}{\sigma_1}, \quad (1.93')$$

и, следовательно,

$$m_1\{n \mid \sigma_0^2\} = 2\,\frac{(1-\alpha)\ln\dfrac{1-\alpha}{\beta} + \alpha\ln\dfrac{\alpha}{1-\beta}}{2\ln\dfrac{\sigma_1}{\sigma_0} - 1 + \dfrac{\sigma_0^2}{\sigma_1^2}}, \quad (1.94)$$

$$m_1\{n \mid \sigma_1^2\} = 2\,\frac{\beta\ln\dfrac{\beta}{1-\alpha} + (1-\beta)\ln\dfrac{1-\beta}{\alpha}}{\dfrac{\sigma_1^2}{\sigma_0^2} - 1 - 2\ln\dfrac{\sigma_1}{\sigma_0}}. \quad (1.94')$$

Из (1.94) и (1.94′) видно, что средний размер выборки, необходимый для принятия окончательного решения, увеличивается по мере приближения отношения $\frac{\sigma_1}{\sigma_0}$ к единице.

1.3.3. Параметр экспоненциального распределения. Пусть $x_1, \ldots, x_n$ представляют независимые выборочные значения, принадлежащие *экспоненциальному* распределению (см. первую книгу, стр. 106)

$$w_1(x) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0. \quad (1.95)$$

Проверяется простая гипотеза H_0 о том, что $\lambda = \lambda_0$, против простой альтернативы H_1, что параметр распределения $\lambda = \lambda_1 > \lambda_0$. Логарифм отношения правдоподобия в этом случае равен

$$\ln l(x_1, \ldots, x_n) = \sum_{k=1}^{n} \ln\frac{\lambda_1 e^{-\lambda_1 x_k}}{\lambda_0 e^{-\lambda_0 x_k}} = n\ln\frac{\lambda_1}{\lambda_0} - (\lambda_1 - \lambda_0)\sum_{k=1}^{n} x_k. \quad (1.96)$$

Правило выбора решения (для фиксированного заранее размера выборки n) формулируется следующим образом: принимается решение γ_1 (параметр распределения равен λ_1), если

$$\frac{1}{n}\sum_{k=1}^{n} x_k \leqslant \frac{1}{\lambda_1-\lambda_0}\ln\frac{\lambda_1}{\lambda_0 c^{1/n}} = K, \qquad (1.97)$$

и принимается решение γ_0 (параметр равен λ_0), если выполняется неравенство, противоположное (1.97).

Таким образом, процедура проверки гипотезы о параметре экспоненциального распределения вероятностей сводится к сравнению *среднего арифметического* выборочных значений с порогом

$$K = \frac{1}{\lambda_1-\lambda_0}\ln\frac{\lambda_1}{\lambda_0 c^{1/n}}, \qquad (1.97')$$

где c определяется в зависимости от выбранного критерия качества правила решения.

Так же как и при проверке гипотезы о среднем нормальной случайной величины, поверхность, разделяющая допустимую и критическую области, представляет гиперплоскость, перпендикулярную единичному вектору и расположенную от начала координат на расстоянии nK, определяемом по формуле (1.97'). В отличие от (1.67), критическая область (1.97) расположена под гиперплоскостью.

Найдем условные вероятности ошибок первого и второго рода. Известно, что сумма n независимых экспоненциально распределенных случайных величин имеет χ^2-распределение с $2n$ степенями свободы (см. первую книгу, задача 3.16, а также предыдущий параграф). Учитывая необходимую нормировку, находим, что в рассматриваемом примере случайная величина $2\lambda\sum_{k=1}^{n} x_k$ распределена по закону χ^2 с $2n$ степенями свободы. Поэтому

$$\alpha = P\left\{2\lambda_0\sum_{k=1}^{n} x_k \leqslant 2n\lambda_0 K \mid H_0\right\} = \frac{\Gamma(n,\, 2n\lambda_0 K)}{\Gamma(n)}, \qquad (1.98)$$

$$\beta = P\left\{2\lambda_1\sum_{k=1}^{n} x_k > 2n\lambda_1 K \mid H_1\right\} = 1 - \frac{\Gamma(n,\, 2n\lambda_1 K)}{\Gamma(n)}. \qquad (1.98')$$

Величина K в (1.98) и (1.98′) для первых трех критериев, указанных в табл. 1 (см. стр. 40), определяется непосредственно из (1.97′). Для критерия Неймана — Пирсона при заданном α из (1.98) следует [ср. с (1.89)]:

$$K = \frac{1}{2n\lambda_0} \chi^2_{1-\alpha}. \tag{1.99}$$

Здесь $\chi^2_{1-\alpha}$ — процентная точка χ^2-распределения с $2n$ степенями свободы. Порог, определяемый согласно (1.99), не зависит от величины λ_1.

Если $n \gg 1$, то можно, использовав асимптотическую нормальность χ^2-распределения, переписать (1.98) и (1.98′) в виде [см. также (1.49) и (1.50)]

$$\alpha \sim F\left(2\sqrt{2n\lambda_0 K} - 2\sqrt{n}\right), \tag{1.100}$$

$$\beta \sim 1 - F\left(2\sqrt{2n\lambda_1 K} - 2\sqrt{n}\right). \tag{1.100′}$$

При $n \to \infty$ в соответствии с общим результатом, приведенным в § 1.2.5, вероятности ошибок стремятся к нулю, так как при $\lambda_1 > \lambda_0$

$$2\lambda_0 K \sim \frac{1}{\frac{\lambda_1}{\lambda_0} - 1} \ln \frac{\lambda_1}{\lambda_0} < 1,$$

$$2\lambda_1 K \sim \frac{\frac{\lambda_1}{\lambda_0}}{\frac{\lambda_1}{\lambda_0} - 1} \ln \frac{\lambda_1}{\lambda_0} > 1.$$

Для *минимаксного* правила решения при $\Pi_{00} = \Pi_{11} = 0$, $\Pi_{10} = \nu \Pi_{01}$ уравнение, определяющее наименее благоприятную величину априорной вероятности $q_{\text{мм}}$, может быть в соответствии с (1.51), (1.98) и (1.98′) представлено в виде

$$\Gamma(n) - \Gamma\left\{n, \frac{2n\lambda_1}{\lambda_1 - \lambda_0} \ln\left[\frac{\lambda_1}{\lambda_0}\left(\frac{\nu p}{q}\right)^{1/n}\right]\right\} = \\ = \nu\Gamma\left\{n, \frac{2n\lambda_0}{\lambda_1 - \lambda_0} \ln\left[\frac{\lambda_1}{\lambda_0}\left(\frac{\nu p}{q}\right)^{1/n}\right]\right\}, \quad p = 1 - q. \tag{1.101}$$

Переходим к *последовательному* правилу выбора решений. В рассматриваемом примере промежуточная область $G_{\text{пр}}$ определяется при заданных α и β согласно (1.59) и (1.96) следующими неравенствами:

$$\frac{1}{\lambda_1 - \lambda_0} \ln\left[\frac{\lambda_1}{\lambda_0}\left(\frac{\alpha}{1-\beta}\right)^{1/n}\right] < \frac{1}{n}\sum_{k=1}^{n} x_k < \\ < \frac{1}{\lambda_1 - \lambda_0} \ln\left[\frac{\lambda_1}{\lambda_0}\left(\frac{1-\alpha}{\beta}\right)^{1/n}\right]. \tag{1.102}$$

В том случае, когда нарушается левое неравенство, т. е.

$$\frac{1}{n}\sum_{k=1}^{n} x_k \leqslant \frac{1}{\lambda_1-\lambda_0}\ln\left[\frac{\lambda_1}{\lambda_0}\left(\frac{\alpha}{1-\beta}\right)^{1/n}\right], \qquad (1.102')$$

принимается решение γ_1: параметр распределения равен λ_1 [ср. с (1.97)], а если нарушается правое неравенство, т. е.

$$\frac{1}{n}\sum_{k=1}^{n} x_k \geqslant \frac{1}{\lambda_1-\lambda_0}\ln\left[\frac{\lambda_1}{\lambda_0}\left(\frac{1-\alpha}{\beta}\right)^{1/n}\right], \qquad (1.102'')$$

то принимается решение γ_0: параметр распределения равен λ_0.

Определим среднее значение размера выборки, необходимого для принятия окончательного решения. Из (1.63), (1.63′), учитывая, что среднее значение случайной величины, распределенной по экспоненциальному закону (1.95), равно $\frac{1}{\lambda}$ (см. стр. 97 в первой книге), имеем

$$m_{10}=m_1\{\ln l(x)\mid\lambda_0\}=$$
$$\ln\frac{\lambda_1}{\lambda_0}-(\lambda_1-\lambda_0)\frac{1}{\lambda_0}=1-\frac{\lambda_1}{\lambda_0}+\ln\frac{\lambda_1}{\lambda_0}, \qquad (1.103)$$
$$m_{11}=m_1\{\ln l(x)\mid\lambda_1\}=$$
$$=\ln\frac{\lambda_1}{\lambda_0}-(\lambda_1-\lambda_0)\frac{1}{\lambda_1}=\frac{\lambda_0}{\lambda_1}-1+\ln\frac{\lambda_1}{\lambda_0}. \qquad (1.103')$$

Следовательно, при $\lambda_1>\lambda_0$

$$m_1\{n\mid\lambda_0\}=\frac{(1-\alpha)\ln\frac{1-\alpha}{\beta}+\alpha\ln\frac{\alpha}{1-\beta}}{\frac{\lambda_1}{\lambda_0}-1-\ln\frac{\lambda_1}{\lambda_0}}, \qquad (1.104)$$

$$m_1\{n\mid\lambda_1\}=\frac{\beta\ln\frac{\beta}{1-\alpha}+(1-\beta)\ln\frac{1-\beta}{\alpha}}{\frac{\lambda_0}{\lambda_1}-1+\ln\frac{\lambda_1}{\lambda_0}}. \qquad (1.104')$$

1.4. СЛОЖНЫЕ ГИПОТЕЗЫ

1.4.1. Классификация задач двоичного выбора. В предыдущем разделе было подробно рассмотрено решение задачи проверки простой гипотезы против простой альтернативы. Оказалось, что для всех принятых в начале главы (см. § 1.1.6) критериев качества (байесовского, минимакс-

ного, Неймана — Пирсона, последовательного) решение может быть доведено до замкнутых формул или очень простых асимптотических соотношений (справедливых при большом размере выборки). Однако указанная задача не всегда отражает действительное положение вещей. Часто и проверяемая гипотеза, и альтернатива, каждая порознь или обе вместе, могут быть *сложными*. Тогда возникает проблема выбора одного из двух решений, γ_H или γ_K, относительно истинности гипотезы H или ее альтернативы K в условиях, когда H и K представляют, вообще говоря, множества гипотез (дискретные или континуальные).

Ограничимся в этом разделе изложением параметрической теории, в которой рассматриваются процедуры проверки гипотез о принадлежности выборки к распределению *данного класса*, определяемого одним или несколькими *параметрами*. Точные значения этих параметров (или некоторых из них) неизвестны. Априорная информация о любом параметре s содержит лишь указание на то, что он принадлежит совокупности возможных значений интервала параметров (или пространства параметров, если речь идет о нескольких параметрах s, u, v, ...). Таким образом, пространство возможных состояний трактуется в рассматриваемом случае как пространство параметров заданного класса распределений вероятностей.

Далее будут рассмотрены задачи двух видов: 1) проверка простой гипотезы $s = s_0$ против сложной альтернативы $s \in S_K$, 2) проверка сложной гипотезы $s \in S_H$ против сложной альтернативы $s \in S_K$.

Относительно множеств S_H и S_K возможных значений параметра s могут вводиться различные предположения. Например, сложная гипотеза может быть односторонней вида $s > s_0$. Множества (интервалы) S_H и S_K могут не пересекаться, а могут иметь и общие элементы. Ж

1.4.2. Байесовское решение. Рассмотрим байесовское решение для общего случая проверки сложной гипотезы против сложной альтернативы. Пусть q и $p = 1 - q$ — априорные вероятности того, что неизвестный параметр s принадлежит непересекающимся множествам S_H и S_K соответственно, и пусть $w_{1H}(s)$ и $w_{1K}(s)$ — плотности вероятности этого параметра на указанных множествах. Тогда априорное распределение параметра s можно представить в виде

$$W_1(s) = qw_{1H}(s) + pw_{1K}(s). \qquad (1.105)$$

Так как множества S_H и S_K могут быть дискретными, непрерывными или смешанными, то соответственно $w_{1H}(s)$ и $w_{1K}(s)$ будут представляться суммами дельта-функций, непрерывными функциями или плотностями смешанного типа (см. первую книгу стр. 49). Например, если проверяется простая гипотеза, что $s = s_0$, против сложной альтернативы, то в (1.105) в первом слагаемом $w_{1H}(s) = \delta(s - s_0)$. (Заметим, что точка $s = s_0$ может в этом случае принадлежать множеству K, если только плотность вероятности $w_{1K}(s)$ в этой точке непрерывная.)

Так же как и в § 1.1.2, два возможных решения γ_0 и γ_1 связаны с принятием или отклонением гипотезы H, и поэтому сохраняется платежная матрица (1.15). Однако в отличие от § 1.2.2 условные риски [см. (1.17) и (1.17′)] становятся *функциями* неизвестного параметра s:

$$r_0(s) = \Pi_{00}[1 - \alpha(s)] + \Pi_{01}\alpha(s), \tag{1.106}$$

$$r_1(s) = \Pi_{10}\beta(s) + \Pi_{11}[1 - \beta(s)], \tag{1.106′}$$

где

$$\alpha(s) = \int_{G_1}\ldots\int W_n(x_1, \ldots, x_n | s)\,dx_1 \ldots dx_n, \quad s \in S_H, \tag{1.107}$$

$$\beta(s) = \int_{G_0}\ldots\int W_n(x_1, \ldots, x_n | s)\,dx_1 \ldots dx_n, \quad s \in S_K. \tag{1.107′}$$

Средняя функция риска записывается теперь в виде [см. (1.16)]

$$R = q\int_{S_H} r_0(s)\,w_{1H}(s)\,ds + p\int_{S_K} r_1(s)\,w_{1K}(s)\,ds =$$

$$= q\left\{\Pi_{00}\left[1 - \int_{S_H} w_{1H}(s)\,\alpha(s)\,ds\right] + \Pi_{01}\int_{S_H} w_{1H}(s)\,\alpha(s)\,ds\right\} +$$

$$+ p\left\{\Pi_{10}\int_{S_K} w_{1K}(s)\,\beta(s)\,ds + \Pi_{11}\left[1 - \int_{S_K} w_{1K}(s)\,\beta(s)\,ds\right]\right\} =$$

$$= q\Pi_{00} + p\Pi_{10} + q(\Pi_{01} - \Pi_{00})\int_{S_H} w_{1H}(s)\,\alpha(s)\,ds -$$

$$- p(\Pi_{10} - \Pi_{11})\left[1 - \int_{S_K} w_{1K}(s)\,\beta(s)\,ds\right].$$

Подставляя вместо α и β их выражения из (1.107) и (1.107′), получим

$$R = q\Pi_{00} + p\Pi_{10} - \int_{G_1}\dots\int [p(\Pi_{10} - \Pi_{11}) \times$$

$$\times \int_{S_K} w_{1K}(s) W_n(x_1, \dots, x_n | s)\, ds - q(\Pi_{01} - \Pi_{00}) \times$$

$$\times \int_{S_H} w_{1H}(s) W_n \times (x_1, \dots, x_n | s)\, ds]\, dx_1 \dots dx_n. \qquad (1.108)$$

Из (1.108) те же соображения, которые были указаны в § 1.2.3, приводят к следующему оптимальному правилу, основанному на минимизации среднего риска: принимается решение γ_1 (отвергается гипотеза H), если для наблюдаемой выборки

$$\frac{p \int_{S_K} w_{1K}(s) W_n(x_1, \dots, x_n | s)\, ds}{q \int_{S_H} w_{1H}(s) W_n(x_1, \dots, x_n | s)\, ds} \geqslant \frac{\Pi_{01} - \Pi_{00}}{\Pi_{10} - \Pi_{11}} = c^*, \qquad (1.109)$$

и принимается решение γ_0 (утверждается истинность гипотезы H), если выполняется неравенство, противоположное (1.109).

Правило (1.20) является частным случаем (1.109) при

$$w_{1H}(s) = \delta(s - s_0), \quad w_{1K}(s) = \delta(s - s_1).$$

При проверке простой гипотезы против сложной альтернативы следует пользоваться частным случаем (1.109) при $w_{1H}(s) = \delta(s - s_0)$, т. е. в знаменателе вместо интеграла следует подставить функцию правдоподобия $W_n(x_1, \dots$ $\dots, x_n | s_0)$. Условные вероятности $\alpha(s)$ и $\beta(s)$ ошибок первого и второго рода для байесовского правила (1.109) определяются по формулам (1.25) и (1.26), в которых теперь F_{10} и F_{11} — интегральные функции распределения отношения усредненных функций правдоподобия

$$\Lambda(x_1, \dots, x_n) = \frac{\int_{S_H} w_{1K}(s) W_n(x_1, \dots, x_n | s)\, ds}{\int_{S_H} w_{1H}(s) W_n(x_1, \dots, x_n | s)\, ds}, \qquad (1.110)$$

вычисленные в предположении, что параметр s распределения выборки $(x_1, \dots, x_n)$ принадлежит множеству S_H и S_K соответственно.

Если гипотеза H — простая, то

$$w_{1H}(s) = \delta(s - s_0)$$

и

$$\Lambda(x_1, \ldots, x_n) = \int\limits_{S_K} w_{1K}(s)\, l(x_1, \ldots, x_n \mid s)\, ds, \quad (1.110')$$

т. е. равна усредненному отношению правдоподобия по случайному параметру.

Полные условные вероятности ошибок первого и второго рода (взвешенные с плотностью вероятности появления отдельного значения параметра s) равны

$$\alpha = \int\limits_{S_H} w_{1H}(s)\, \alpha(s)\, ds, \quad (1.111)$$

$$\beta = \int\limits_{S_K} w_{1K}(s)\, \beta(s)\, ds. \quad (1.111')$$

Если области S_H и S_K *перекрываются*, то уже невозможно достоверно установить, относится ли данное значение параметра к области проверяемой гипотезы или к области альтернативы. В связи с этим понятие платы за правильное или ложнсе решение теряет смысл, так как решение может быть и правильным, и ложным, если s принадлежит пересечению областей S_H и S_K. Функция потерь, сопоставляющая каждой комбинации $s \in S_H$, $s \in S_K$ и решений γ_0, γ_1 плату Π_{jk}, теперь зависит от апостериорных вероятностей того, что $s \in S_H$ или $s \in S_K$, если значение параметра равно s (при непересекающихся областях эти вероятности равны единице и нулю). Одним из возможных подходов при этом будет следующий. Пусть $\Pi_{00} = \Pi_{11} = 0$ и предположим, что плата за ошибочное решение пропорциональна апостериорной вероятности того, что данное s принадлежит истинной области, т. е.

$$\Pi_{10}(s) = \Pi_{10} P\{s \in S_K \mid s\} = \Pi(S_K, \gamma_0), \quad (1.112)$$

$$\Pi_{01}(s) = \Pi_{01} P\{s \in S_H \mid s\} = \Pi(S_H, \gamma_1), \quad (1.112')$$

где

$$P\{s \in S_K \mid s\} = \frac{P\{s \in S_K\}\, P\{s \mid s \in S_K\}}{P\{s \in S_K\}\, P\{s \mid s \in S_K\} + P\{s \in S_H\}\, P\{s \mid s \in S_H\}} =$$

$$= \frac{p w_{1K}(s)}{p w_{1K}(s) + q w_{1H}(s)}, \quad (1.113)$$

$$P\{s\in S_H\mid s\}=\frac{P\{s\in S_H\}P\{s\mid s\in S_H\}}{P\{s\in S_K\}P\{s\mid s\in S_K\}+P\{s\in S_H\}P\{s\mid s\in S_H\}}=$$
$$=\frac{qw_{1H}(s)}{pw_{1K}(s)+qw_{1H}(s)}. \qquad (1.113')$$

Если области S_H и S_K не перекрываются, то $P\{s\in S_K\mid s\}=1$, $P\{s\in S_H\mid s\}=1$ и $\Pi_{10}(s)=\Pi_{10}$, $\Pi_{01}(s)=\Pi_{01}$, т. е. приходим к предыдущему случаю.

Выражение для условных рисков теперь преобразуются к виду

$$r_0(s)=\Pi_{01}(s)\,\alpha(s),$$
$$r_1(s)=\Pi_{10}(s)\,\beta(s),$$

а средний риск находится по формуле

$$R=q\int\limits_{S_H}\Pi_{01}(s)\,\alpha(s)\,w_{1H}(s)\,ds+p\int\limits_{S_K}\Pi_{10}(s)\,\beta(s)\,w_{1K}(s)\,ds.$$

1.4.3. Максимум апостериорной вероятности и максимальное правдоподобие. Апостериорные вероятности того, что неизвестный параметр s принадлежит S_H или S_K, когда наблюдается выборка $(x_1,\ldots,x_n)$, могут быть записаны в виде

$$P\{s\in S_H\mid x_1,\ldots,x_n\}=\frac{q\int\limits_{S_H}w_{1H}(s)\,W_n(x_1,\ldots,x_n\mid s)\,ds}{q\int\limits_{S_H}w_{1H}(s)\,W_n(x_1,\ldots,x_n\mid s)\,ds+p\int\limits_{S_K}w_{1K}(s)\,W_n(x_1,\ldots,x_n\mid s)\,ds}, \qquad (1.114)$$

$$P\{s\in S_K\mid x_1,\ldots,x_n\}=\frac{p\int\limits_{S_K}w_{1K}(s)\,W_n(x_1,\ldots,x_n\mid s)\,ds}{q\int\limits_{S_H}w_{1H}(s)\,W_n(x_1,\ldots,x_n\mid s)\,ds+p\int\limits_{S_K}w_{1K}(s)\,W_n(x_1,\ldots,x_n\mid s)\,ds}. \qquad (1.114')$$

Из того условия, что для отклонения гипотезы H первая из приведенных вероятностей должна быть не больше второй, находим правило выбора решения

$$\frac{p\int\limits_{S_K}w_{1K}(s)\,W_n(x_1,\ldots,x_n\mid s)\,ds}{q\int\limits_{S_H}w_{1H}(s)\,W_n(x_1,\ldots,x_n\mid s)\,ds}\geqslant 1, \qquad (1.115)$$

которое является частным случаем байесовского правила (1.109), когда $\Pi_{01} - \Pi_{00} = \Pi_{10} - \Pi_{11}$.

Критерий максимального правдоподобия не основывается в рассматриваемом случае на строгих соображениях оптимальности. Одно из возможных правил, соответствующих интуитивному представлению о максимальном правдоподобии, может быть представлено в виде: принимается решение γ_1 (отвергается H), если для наблюдаемой выборки выполняется неравенство

$$\frac{\max\limits_{s \in S_K} W_n(x_1, \ldots, x_n \mid s)}{\max\limits_{s \in S_H} W_n(x_1, \ldots, x_n \mid s)} \geqslant 1, \tag{1.116}$$

и принимается решение γ_0, если выполняется неравенство, противоположное (1.116). Заметим, что теперь уже правило максимального правдоподобия *не является частным случаем байесовского*, как это было при проверке простой гипотезы против простой альтернативы.

1.4.4. Критерий Неймана — Пирсона. Переходим теперь к критерию Неймана — Пирсона. Как было установлено в § 1.2.4, при проверке простой гипотезы против простой альтернативы всегда существует правило разбиения пространства выборок на критическую и допустимую области, которое минимизирует вероятность ошибки второго рода (имеет максимальную мощность) среди всех других правил, для которых вероятность ошибки первого рода (уровень значимости) не превосходит заданной величины α. Когда гипотеза H простая, а альтернатива K сложная, то $\beta = \beta(s)$ и можно попытаться найти такое правило выбора решения (т. е. разбиение пространства выборок на две области G_0 и G_1), которое для заданной вероятности α ошибок первого рода минимизирует вероятность ошибки второго рода $\beta(s)$ [или максимизирует мощность $1 - \beta(s)$] для *всех* простых альтернатив, содержащихся в сложной альтернативе K. Такое правило называют *равномерно наиболее мощным*. Если существует равномерно наиболее мощное правило выбора решения при проверке простой гипотезы против сложной альтернативы, то оно по существу не отличается от такого же правила, соответствующего *простой* альтернативе, так как при этом неоднозначность, возникающая из-за того, что S_K представляет множество значений параметра s, не имеет значения (так как критическая область не зависит от $s \in S_K$).

Существование равномерно наиболее мощного правила выбора решения является скорее исключением, нежели правилом. Если такое правило при решении конкретной задачи отсутствует, то можно попытаться сузить класс правил и искать в этом меньшем классе правил равномерно наиболее мощное. К одному из таких суженных классов принадлежат так называемые *несмещенные* правила. Эти правила должны удовлетворять следующему условию: вероятность отвергнуть ложную гипотезу не меньше вероятности отвергнуть правильную *). Иначе говоря, вероятность α ошибки первого рода является нижней границей значений функции мощности $1-\beta(s)$ для всех значений s, т. е.

$$1-\beta(s)\geqslant\alpha. \tag{1.117}$$

Если $\beta(s)$ — непрерывная функция, то минимальное значение $1-\beta(s)$ достигается при $s=s_0$ и в точности равно α, так как

$$1-\beta(s_0)=P\{(x_1, \ldots, x_n)\in G_1\,|\,s_0\}=\alpha. \tag{1.118}$$

Равномерно наиболее мощное правило, если оно существует, всегда является несмещенным. Если же такого правила не существует, то все же может быть найдено несмещенное равномерно наиболее мощное правило.

На рис. 7 показаны кривые мощностей, иллюстрирующие понятие равномерно наиболее мощного правила, несмещенного равномерно наиболее мощного правила и смещенного правила выбора решения при проверке простой гипотезы против сложной альтернативы.

Заметим, что оптимальное по критерию Неймана — Пирсона правило выбора решения при проверке простой гипотезы против сложной альтернативы не является, вообще говоря, частным случаем байесовского правила, как это имело место при простой альтернативе (см. стр. 36).

Приведенные выше понятия обобщаются и тогда, когда проверяемая гипотеза H сложная. В этом случае вероятность ошибки первого рода (уровень значимости) также зависит от параметра s, принадлежащего некоторому множеству S_H. Однако иногда можно указать совокупность правил выбора решения (совокупность критических областей), которой соответствует постоянный для всех $s\in S_H$ уровень значи-

---

*) Заметим, что в примере, рассмотренном в § 1.1.2, условие (1.5) — простейшее условие несмещенности.

мости α. Теперь задача состоит в том, чтобы из этой совокупности правил отыскать такое, которое является равномерно наиболее мощным для всех значений $s \in S_K$. Конечно, как и в рассматриваемой выше более простой ситуации, решение указанной задачи может и не существовать. Тогда следует попытаться найти его в более узком классе правил,

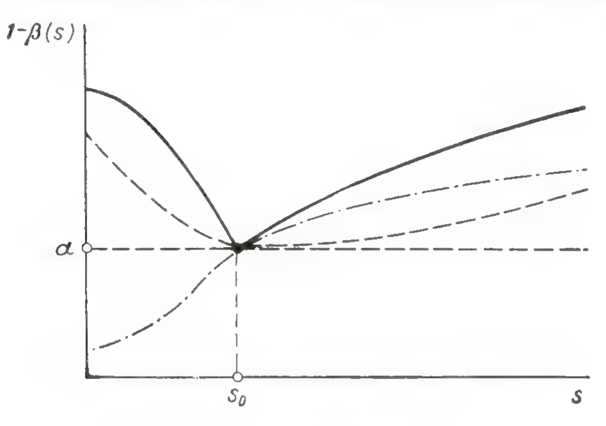


Рис. 7. Функции мощности:

——— равномерно наиболее мощное правило, — — — несмещенное равномерно наиболее мощное правило, —· ·· смещенное правило.

вводя дополнительные предположения. Одним из таких предположений является несмещенность, т. е. условие, что

$$P\{(x_1, \ldots, x_n) \in G_1 \mid s \in S_H\} \leqslant P\{(x_1, \ldots, x_n) \in G_1 \mid s \in S_K\}. \tag{1.119}$$

1.4.5. Минимаксное правило. Если плотности $w_{1H}(s)$ и $w_{1K}(s)$ известны, а априорные вероятности q и $p = 1 - q$ неизвестны, то можно усреднить условные функции риска

$$r_0 = \int_{S_H} r_0(s)\, w_{1H}(s)\, ds, \tag{1.120}$$

$$r_1 = \int_{S_K} r_1(s)\, w_{1K}(s)\, ds \tag{1.120$'$}$$

и находить минимаксное правило, приравнивая величины усредненных средних рисков, аналогично тому, как это делалось в § 1.2.6 для нахождения минимаксного правила

при проверке простой гипотезы против простой альтернативы.

В том случае, когда известно q, но неизвестны $w_{1H}(s)$ и $w_{1K}(s)$, решение задачи об определении минимаксного правила может и не быть однозначным, а если и априорная вероятность q неизвестна, необходимо выдвижение дополнительных критериев качества.

1.4.6. Последовательный анализ. К сожалению, для последовательного анализа нет ситуации, аналогичной той,

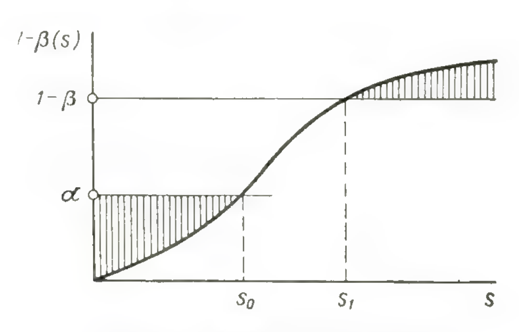


Рис. 8. Функция мощности.

которая имела место при использовании критерия Неймана — Пирсона, когда одно и то же правило могло быть наилучшим для любого значения параметра s, принадлежащего области альтернативы. Последовательное правило, минимизирующее средние размеры выборки для одного значения s, не является оптимальным в этом смысле при другом значении параметра.

Одна из возможных постановок задач в этом случае заключается в следующем. Проверяется односторонняя сложная гипотеза H о том, что $s \leqslant s_0$, против сложной альтернативы $s > s_0$. Выдвигается требование к качеству правила выбора решения, состоящее в том, чтобы

$$1 - \beta(s) \leqslant \alpha, \quad s \leqslant s_0, \tag{1.121}$$

$$1 - \beta(s) \geqslant 1 - \beta, \quad s \geqslant s_1, \quad s_1 > s_0. \tag{1.121'}$$

Условия (1.121), (1.121′) выполняются, если

$$1 - \beta(s_0) = \alpha, \quad \beta(s_1) = \beta \tag{1.122}$$

и вероятность $\beta(s)$ ошибки второго рода — невозрастающая функция параметра s (рис. 8). Из оптимального свой-

ства последовательного анализа очевидно, что среди всех правил, удовлетворяющих условиям (1.121) и (1.121′), последовательное правило выбора решения, состоящее в сравнении отношения правдоподобия с порогами

$$c_0 = \frac{\beta}{1-\alpha}, \quad c_1 = \frac{1-\beta}{\alpha},$$

минимизирует средние размеры выборки при $s = s_0$ и $s = s_1$. Обычно функция $m_1\{n \mid s\}$ имеет максимум между s_0 и s_1 и убывает, когда s удаляется от точки максимума в любом направлении. Может оказаться, что этот максимум меньше наименьшего фиксированного размера выборки, при котором существует (непоследовательное) правило, удовлетворяющее условиям (1.121) и (1.121′). Однако это бывает не всегда, как указывает пример, приведенный в [7].

Точное определение функции мощности $1-\beta(s)$ и среднего размера выборки $m_1\{n \mid s\}$ для последовательного анализа представляет трудную задачу. Если трансцендентное уравнение

$$m_1\left\{\left[\frac{w_1(x \mid s_1)}{w_1(x \mid s_0)}\right]^h \middle| s\right\} = 1 \qquad (1.123)$$

имеет решение $h = h(s)$, то с некоторым приближением

$$1-\beta(s) \approx \frac{1-c_0^{h(s)}}{c_1^{h(s)}-c_0^{h(s)}} \qquad (1.124)$$

и

$$m_1\{n \mid s\} \approx \frac{[1-\beta(s)]\ln c_1 + \beta(s)\ln c_0}{m_1\left\{\ln\frac{w_1(x \mid s_1)}{w_1(x \mid s_0)} \middle| s\right\}} \qquad (1.125)$$

в предположении, что знаменатель в (1.125) отличен от нуля Если последнее условие не выполняется, то вместо (1.125) следует использовать другое соотношение, указанное в [2, стр. 223], в котором вместо среднего в знаменателе появляется дисперсия логарифма одномерного отношения правдоподобия:

$$m_1\{n \mid s\} \approx \frac{-\ln c_0 \ln c_1}{m_2\left\{\ln\frac{w_1(x \mid s_1)}{w_1(x \mid s_0)} \middle| s\right\}}. \qquad (1.126)$$

1.4.7. Проверка сложных гипотез о среднем нормальной случайной величины. Проиллюстрируем изложенные в этом разделе подходы на примерах проверки гипотез о параметрах

нормального распределения. Начнем с проверки простой гипотезы H о том, что среднее значение нормальной случайной величины равно a_0, против сложной альтернативы K, состоящей в том, что среднее — любое действительное число, не равное a_0. Предполагается, что дисперсия σ^2 нормальной случайной величины известна точно и что выборка, на основании которой проверяется гипотеза, состоит из независимых элементов.

Для получения байесовского правила решения необходимо задать величины c^*, $\mu = q/p$, а также априорное распределение среднего, когда справедлива альтернатива K. Пусть это распределение также нормальное с параметрами (a_1, σ_1^2). Тогда из (1.110) при $w_{1H}(s) = \delta(s - a_0)$ получаем следующее выражение функции $\Lambda(x_1, \ldots, x_n)$:

$$\Lambda(x_1, \ldots, x_n) = \exp\left[\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k - a_0)^2\right]\frac{1}{\sigma_1\sqrt{2\pi}} \times$$

$$\times \int_{-\infty}^{\infty} e^{-\frac{(s-a_1)^2}{2\sigma_1^2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-s)^2}\, ds = \left(1 + \frac{n\sigma_1^2}{\sigma^2}\right)^{-\frac{1}{2}} \times$$

$$\times \exp\left\{\frac{1}{2}\left[\frac{2(a_1 - a_0)}{\sigma^2}\sum_{i=1}^{n} x_i - \frac{n(a_1^2 - a_0^2)}{\sigma^2} + \right.\right.$$

$$\left.\left. + \frac{\left(\frac{\sigma_1}{\sigma}\right)^2}{1 + n\left(\frac{\sigma_1}{\sigma}\right)^2}\left(\sum_{k=1}^{n}\frac{x_k - a_1}{\sigma}\right)^2\right]\right\},$$

откуда следует байесовское правило выбора решения: отвергается простая гипотеза H о том, что среднее значение равно a_0, если при $a_1 > a_0$

$$\frac{1}{n}\sum_{i=1}^{n} x_i + \frac{n\left(\frac{\sigma_1}{\sigma}\right)^2}{2\left(1 + \frac{n\sigma_1^2}{\sigma^2}\right)\frac{a_1 - a_0}{\sigma^2}}\left(\frac{1}{n}\sum_{k=1}^{n}\frac{x_k - a_1}{\sigma}\right)^2 >$$

$$> \frac{a_0 + a_1}{2} + \frac{\sigma^2}{n(a_1 - a_0)}\ln\left[c\left(1 + \frac{n\sigma_1^2}{\sigma^2}\right)\right] = K, \qquad (1.127)$$

где $c = \mu c^*$. При $c = \mu$ получаем правило выбора решения, соответствующее максимуму апостериорной вероятности.

Рассмотрим теперь критерий Неймана — Пирсона. Как было показано в § 1.3.1, оптимальное по этому критерию правило выбора решения при проверке простой гипотезы о величине среднего против простой альтернативы, определяемое порогом (1.76), *не зависит* от альтернативы (при условии $s > a_0$ или $s < a_0$). Поэтому правило, согласно которому отвергается гипотеза H, если

$$\frac{1}{n}\sum_{k=1}^{n} x_k > a_0 + \frac{\sigma}{\sqrt{n}}\, x_\alpha, \tag{1.128}$$

является равномерно наиболее мощным относительно сложной альтернативы, для которой $s > a_0$. Если $s < a_0$, равномерно наиболее мощное правило при сложной альтернативе определяет критическую область в виде неравенства

$$\frac{1}{n}\sum_{k=1}^{n} x_k < a_0 - \frac{\sigma}{\sqrt{n}}\, x_\alpha. \tag{1.129}$$

Однако если в качестве сложной альтернативы рассматривать все действительные значения, то равномерно наиболее мощного правила не существует. Функция мощности для правила (1.128) согласно (1.77) имеет вид *)

$$1-\beta(s) = 1 - F\left(x_\alpha - \frac{s-a_0}{\sigma}\sqrt{n}\right). \tag{1.130}$$

При $s > a_0$, $1-\beta(s) > \alpha$, функция мощности монотонно возрастает при увеличении s и $1-\beta(a_0) = \alpha$. Но если $s < a_0$, то $1-\beta(s) < \alpha$, причем функция мощности убывает при уменьшении s. Таким образом, правило (1.128) *смещенное*.

Можно показать [6], что несмещенное равномерно наиболее мощное правило определяется в этом случае критической областью

$$\left|\frac{1}{n}\sum_{k=1}^{n} x_i - a_0\right| > \frac{x_{\frac{\alpha}{2}}\sigma}{\sqrt{n}}. \tag{1.131}$$

---

*) Заметим, что при $s = a_1$ формула (1.130) не отличается от (1.77).

Функция мощности, соответствующая правилу (1.131), показана на рис. 9 (пунктирная линия) и имеет вид

$$1-\beta(s)$$
$$1-\left[F\left(x_{\frac{\alpha}{2}}-\frac{s-a_0}{\sigma}\sqrt{n}\right)-F\left(-x_{\frac{\alpha}{2}}-\frac{s-a_0}{\sigma}\sqrt{n}\right)\right]. \qquad (1.132)$$

Функция (1.132) достигает минимума при $s=a_0$ и

$$1-\beta(a_0)=2-2F(x_{\frac{\alpha}{2}})=\alpha.$$

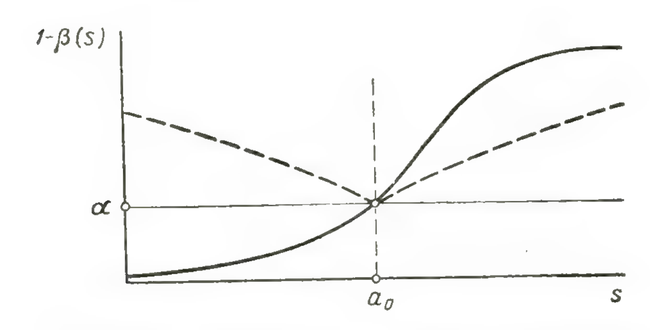


Рис. 9. Функции мощности при проверке гипотезы о среднем:

——— смещенное правило, — — — несмещенное равномерно наиболее мощное правило.

При всех $s>a_0$ функция (1.130) превышает (1.132), так как для произвольного $c>0$

$$F(x_{\frac{\alpha}{2}}-c)-F(x_\alpha-c)>F(-x_{\frac{\alpha}{2}}-c).$$

При $s=a_0$ обе функции совпадают, но при $s<a_0$ мощность правила (1.128) *меньше* мощности правила (1.131).

Рассмотрим, наконец, *последовательное* правило проверки односторонней гипотезы о том, что среднее значение нормальной случайной величины $s\leqslant a_0$, против сложной альтернативы $s>a_0$. В соответствии с § 1.4.6 выдвигаем дополнительное условие [см. (1.121) и (1.121′)]: при заданных α и β

$$1-\beta(s)\leqslant\alpha, \quad s\leqslant s_0,$$
$$1-\beta(s)\geqslant 1-\beta, \quad s\geqslant s_1>s_0.$$

Для определения функции мощности $1-\beta(s)$ необходимо решить трансцендентное уравнение (1.123). Левая часть этого уравнения для рассматриваемого примера равна

$$m_1\left\{\exp\left[\frac{h}{2\sigma^2}[(x-s_0)^2-(x-s_1)^2]\right]\Big| s\right\}=$$

$$=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(\frac{s_0^2-s_1^2}{2\sigma^2}h\right)\times$$

$$\times\exp\left\{-\frac{1}{2\sigma^2}[(x-s)^2+2(s_0-s_1)hx]\right\}dx$$

$$=\exp\left\{\frac{s_0^2-s_1^2}{2\sigma^2}h-\frac{s^2}{2\sigma^2}+\frac{1}{2\sigma^2}[s-(s_0-s_1)h]^2\right\}.$$

Теперь в соответствии с (1.123) необходимо это выражение приравнять единице, что равносильно приравниванию нулю показателя экспоненты, т. е.

$$(s_0^2-s_1^2)h-s^2+[s-(s_0-s_1)h]^2=0,$$

откуда находим

$$h(s)=\frac{s_1+s_0-2s}{s_1-s_0}=1-\frac{2(s-s_0)}{s_1-s_0}. \qquad (1.133)$$

Теперь по формуле (1.124) находим

$$1-\beta(s)\approx\frac{1-c_0^{1-2\frac{s-s_0}{s_1-s_0}}}{c_1^{1-2\frac{s-s_0}{s_1-s_0}}-c_0^{1-2\frac{s-s_0}{s_1-s_0}}}, \qquad (1.134)$$

где $c_0=\frac{\beta}{1-\alpha}$; $c_1=\frac{1-\beta}{\alpha}$.

Из (1.133) и (1.134) следует, что

$$h(s_0)=1,\quad 1-\beta(s_0)=\frac{1-c_0}{c_1-c_0}=\alpha, \qquad (1.135)$$

$$h(s_1)=-1,\quad 1-\beta(s_1)=\frac{1-\frac{1}{c_0}}{\frac{1}{c_1}-\frac{1}{c_0}}=1-\beta. \qquad (1.135')$$

Кроме того, функция $1-\beta(s)$ — монотонно возрастающая (рис. 10) от нуля (при $s\to-\infty$) до единицы (при $s\to\infty$). При $s=\frac{s_1+s_0}{2}$

$$1-\beta\left(\frac{s_1+s_0}{2}\right)=\frac{\ln c_0}{\ln\frac{c_0}{c_1}}=\frac{\ln\frac{1-\alpha}{\beta}}{\ln\frac{(1-\alpha)(1-\beta)}{\alpha\beta}}. \qquad (1.136)$$

Для определения зависимости от величины s среднего размера выборки, необходимого для принятия решения, найдем в соответствии с (1.125)

$$m_1\left\{\ln\frac{w_1(x\mid s_1)}{w_1(x\mid s_0)}\,\middle|\, s\right\} = m_1\left\{\frac{s_0^2-s_1^2}{2\sigma^2} - \frac{s_0-s_1}{\sigma^2}x\,\middle|\, s\right\} = $$

$$= \frac{s_1-s_0}{\sigma^2}\left(s-\frac{s_1+s_0}{2}\right)$$

и, следовательно,

$$m_1\{n\mid s\} = \frac{[1-\beta(s)]\ln c_1+\beta(s)\ln c_0}{\frac{s_1-s_0}{\sigma^2}\left(s-\frac{s_1+s_0}{2}\right)}. \qquad (1.137)$$

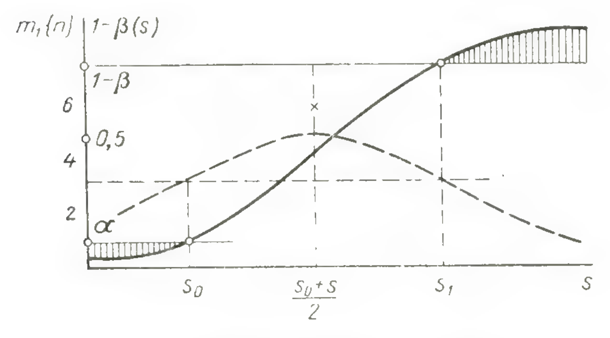


Рис. 10. Последовательное правило при проверке гипотезы о среднем:

——— функция мощности; — — — средний объем выборки, необходимый для принятия решения.

Из (1.137) видно, что при $s\to\pm\infty$ величина $m_1\{n\mid s\}\to$ $\to 0$, а в заданных точках $s=s_0$ и $s=s_1$ [см. (1.135), (1.135')]

$$m_1\{n\mid s_0\} = \frac{\alpha\ln c_1+(1-\alpha)\ln c_0}{-\frac{(s_1-s_0)^2}{2\sigma^2}}, \qquad (1.138)$$

$$m_1\{n\mid s_1\} = \frac{(1-\beta)\ln c_1+\beta\ln c_0}{\frac{(s_1-s_0)^2}{2\sigma^2}}. \qquad (1.139)$$

Эти формулы по существу не отличаются от (1.82) и (1.83).

При $s=\frac{s_1+s_0}{2}$ знаменатель (1.137) обращается в нуль, и в этой точке для расчета $m_1\left\{n\,\middle|\,\frac{s_0+s_1}{2}\right\}$ необходимо

предварительно вычислить второй момент логарифма отношения правдоподобия:

$$m_2\{\ln l(x) \mid s\} = m_1\left\{\frac{(s_1-s_0)^2}{\sigma^4}(x-s)^2 \mid s\right\} - \frac{(s_1-s_0)^2}{\sigma^2}. \quad (1.140)$$

Подставляя (1.140) в (1.126), получаем

$$m_1\left\{n \,\middle|\, \frac{s_0+s_1}{2}\right\} = -\frac{\ln c_0 \ln c_1}{\left(\frac{s_1-s_0}{\sigma}\right)^2}. \quad (1.141)$$

Пунктирная кривая на рис. 10 иллюстрирует характер зависимости среднего размера выборки от s (она построена для $\alpha = \beta = 0{,}1$; см. также табл. 2 на стр. 55). Как видно из этого рисунка, функция $m_1\{n \mid s\}$ имеет максимум при $s = \frac{s_0+s_1}{2}$, который приблизительно на 40% превышает значение минимального среднего размера выборки, необходимого для принятия решения согласно оптимальному последовательному правилу при проверке простой гипотезы $s = s_0$ против простой альтернативы $s = s_1$. Однако этот максимум все же меньше размера выборок при использовании непоследовательного правила с теми же значенияим условных вероятностей ошибок (этот размер выборки отмечен крестиком на рис. 10).

1.4.8. Замечание относительно многоальтернативных задач выбора решения. Ограничимся только байесовским критерием качества *). Как указывалось в § 1.1, в том случае, когда выдвигается $m+1$ простых гипотез H_j ($j = 0, 1, \ldots, m$), то для выбора решения относительно того, которая из них истинная, можно использовать критерий минимума среднего риска [см. (1.8)]. Правило выбора решения состоит в указании способа разбиения пространства выборок на $m+1$ непересекающихся областей и приписывания (для нерандомизированного правила) каждой из областей G_k одного из решений γ_k (верна гипотеза H_k) о том, что имеет место состояние s_k исследуемого явления).

Вероятность того, что по наблюдаемой выборке $x_1, \ldots, x_n$ будет принято решение γ_k, когда в действитель-

*) Исследование многоальтернативных задач начато в последние годы. Некоторые важные примеры многоальтернативных процедур по критерию Неймана — Пирсона содержатся в [7]. О последовательном анализе в многоальтернативных задачах см. [11].

ности верна гипотеза H_j, будет равна

$$P\{\gamma_k | H_j\} \simeq P\{(x_1, \ldots, x_n) \in G_k | s_j\} = \\ = \int_{G_k} \ldots \int W_n(x_1, \ldots, x_n | s_j)\, dx_1 \ldots dx_n. \quad (1.142)$$

Подставляя (1.142) в (1.8), получим выражение для среднего риска

$$R = \sum_{j=0}^{m} \sum_{k=0}^{m} p_j \text{П}_{jk} \int_{G_k} \ldots \int W_n(x_1, \ldots, x_n | s_j)\, dx_1 \ldots dx_n, \quad (1.143)$$

которое зависит от способа разбиения пространства выборок на области G_k, $k = 0, 1, \ldots, m$.

Рассуждая аналогично, как и при выводе неравенства (1.20), можно доказать, что минимальное значение среднего риска R реализует разбиение пространства выборок G, при котором область G_k $(k = 1, \ldots, m)$ определяется *системой* m неравенств:

$$\sum_{i=0}^{m} (\text{П}_{ij} - \text{П}_{ik}) \frac{p_i W_n(x_1, \ldots, x_n | s_i)}{p_0 W_n(x_1, \ldots, x_n | s_0)} \geqslant 0, \quad (1.144)$$

$$j = 0, \ldots, m;\ j \neq k$$

Область G_0 при этом определяется из очевидного условия

$$G_0 = G - \sum_{k=1}^{m} G_k. \quad (1.144')$$

При $m = 1$ система (1.144) содержит одно неравенство, совпадающее с (1.20).

Вводя новые переменные

$$y_i = \frac{p_i}{p_0} l_i(x_1, \ldots, x_n) = \frac{p_i}{p_0} \frac{W_n(x_1, \ldots, x_n | s_i)}{W_n(x_1, \ldots, x_n | s_0)}, \\ i = 1, \ldots, m, \quad (1.145)$$

т. е. отображая точки пространства выборок на m-мерную область отношений правдоподобия, можно систему неравенств записать в виде

$$\sum_{i=1}^{m} (\text{П}_{ij} - \text{П}_{ik}) y_i \geqslant \text{П}_{0k} - \text{П}_{0j}, \qquad j = 0, \ldots, m;\ j \neq k. \quad (1.146)$$

Область, задаваемая системой неравенств (1.146), определяется пересечением гиперплоскостей в m-мерном пространстве.

Проиллюстрируем сказанное выше на простейшем примере многоальтернативной задачи, когда априори известно,

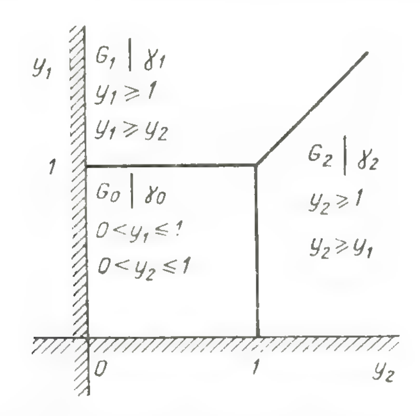


Рис. 11. Области принятия решения.

что возможны три состояния: s_0, s_1, s_2. В соответствии с (1.146) при $m = 2$ байесовское правило решения формулируется так:

1) принимается решение γ_1 о том, что имеет место состояние s_1, если

$$(\Pi_{10} - \Pi_{11})\, y_1 + (\Pi_{20} - \Pi_{21})\, y_2 \geqslant \Pi_{01} - \Pi_{00}, \qquad (1.147)$$

$$(\Pi_{12} - \Pi_{11})\, y_1 + (\Pi_{22} - \Pi_{21})\, y_2 \geqslant \Pi_{01} - \Pi_{02}; \qquad (1.147')$$

2) принимается решение γ_2 о том, что имеет место состояние s_2, если

$$(\Pi_{10} - \Pi_{12})\, y_1 + (\Pi_{20} - \Pi_{22})\, y_2 \geqslant \Pi_{02} - \Pi_{00}, \qquad (1.148)$$

$$(\Pi_{11} - \Pi_{12})\, y_1 + (\Pi_{21} - \Pi_{22})\, y_2 \geqslant \Pi_{02} - \Pi_{01}; \qquad (1.148')$$

3) принимается решение γ_0 о том, что имеет место состояние s_0, если

$$(\Pi_{10} - \Pi_{11})\, y_1 + (\Pi_{20} - \Pi_{21})\, y_2 < \Pi_{01} - \Pi_{00}, \qquad (1.149)$$

$$(\Pi_{10} - \Pi_{12})\, y_1 + (\Pi_{20} - \Pi_{22})\, y_2 < \Pi_{02} - \Pi_{00}). \qquad (1.149')$$

Здесь два отношения правдоподобия

$$y_1 = \frac{p_1}{p_0} \frac{W_n(x_1, \ldots, x_n \mid s_1)}{W_n(x_1, \ldots, x_n \mid s_0)},$$

$$y_2 = \frac{p_2}{p_0} \frac{W_n(x_1, \ldots, x_n \mid s_2)}{W_n(x_1, \ldots, x_n \mid s_0)},$$

$$p_0 = 1 - p_1 - p_2,$$

представляют функциональное преобразование n-мерного случайного вектора с компонентами $(x_1, \ldots, x_n)$ в плоский случайный вектор с неотрицательными компонентами (y_1, y_2). В зависимости от того, в какую из трех непересекающихся областей первой четверти плоскости, определяемых приведенными системами двух неравенств, попадает указанный вектор, принимается одно из трех возможных решений.

На рис. 11 показаны области принятия трех решений для частного случая, когда платы за правильные решения равны нулю, а платы за ошибочные решения равны между собой.

1.5. БОЛЕЕ ОБЩИЕ СЛУЧАИ ВЫБОРА ОДНОГО ИЗ ДВУХ РЕШЕНИЙ

1.5.1. Случай нескольких неизвестных параметров. До сих пор рассматривались правила, связанные с проверкой сложной гипотезы о параметре распределения против сложной альтернативы относительно этого параметра. Результаты могут быть обобщены на случай, когда распределение зависит от многих параметров.

Пусть распределение случайной величины зависит от M параметров $s_1, \ldots, s_M$. Проверяется сложная гипотеза H о том, что указанная совокупность параметров принадлежит области S_H M-мерного пространства, против сложной альтернативы K, что эта совокупность принадлежит области S_K.

Если известны априорные вероятности q и $p = 1 - q$ принадлежности совокупности $s_1, \ldots, s_M$ непересекающимся областям S_H и S_K, а также совместные функции распределения $w_{MH}(s_1, \ldots, s_M)$ и $w_{MK}(s_1, \ldots, s_M)$ этой совокупности на множествах S_H и S_K соответственно и если задана функция потерь, то *байесовским* правилом, обобщающим (1.109), является в этом случае следующее: при-

нимается решение γ_1 (отвергается гипотеза H), если для наблюдаемой выборки

$$\Lambda(x_1, \ldots, x_n) = \frac{\int\limits_{S_K} \ldots \int w_{MK}(s_1, \ldots, s_M) W_n(x_1, \ldots, x_n \mid s_1, \ldots, s_M)\, ds_1 \ldots ds_M}{\int\limits_{S_H} \ldots \int w_{MH}(s_1, \ldots, s_M) W_n(x_1, \ldots, x_n \mid s_1, \ldots, s_M)\, ds_1 \ldots ds_M} > \geqslant \frac{q}{p} c^* - c. \quad (1.150)$$

Полные условные вероятности ошибок первого и второго рода равны

$$\alpha = \int\limits_{S_H} \ldots \int w_{MH}(s_1, \ldots, s_M)\, \alpha(s_1, \ldots, s_M)\, ds_1 \ldots ds_M, \quad (1.151)$$

$$\beta = \int\limits_{S_K} \ldots \int w_{MH}(s_1, \ldots, s_M)\, \beta(s_1, \ldots, s_M)\, ds_1 \ldots ds_M, \quad (1.152)$$

где

$$\alpha(s_1, \ldots, s_M) = P\{\Lambda \geqslant c \mid (s_1, \ldots, s_M) \in S_H\}, \quad (1.153)$$

$$\beta(s_1, \ldots, s_M) = P\{\Lambda < c \mid (s_1, \ldots, s_M) \in S_K). \quad (1.154)$$

Если априорная вероятность p неизвестна, то наименее благоприятная величина этой вероятности $p_{\text{мм}}$ находится из равенства усредненных по областям S_H и S_K условных функций риска [см. § 1.4.5, также (1.106) и (1.106′)]. Подставляя $p_{\text{мм}}$ и $q_{\text{мм}} = 1 - p_{\text{мм}}$ в правую часть неравенства (1.150), приходим к *минимаксному* правилу выбора решения для случая нескольких неизвестных параметров.

Нахождение оптимального правила по критерию Неймана — Пирсона в случае нескольких неизвестных параметров представляет в общем сложную математическую задачу. В конце § 1.4.4 указывалось лишь на возможность обобщения основных понятий (современное состояние этого вопроса изложено в [7]). Как в случае проверки простой гипотезы против простой альтернативы, так и в случае проверки *простой* гипотезы против *сложной* альтернативы оптимизация по критерию Неймана — Пирсона проводилась в классе правил, для которых вероятность ошибки первого

рода (уровень значимости) не превосходила заданной величины α. В некоторых случаях проверки сложной гипотезы против сложной альтернативы можно выделить класс правил (т. е. указать критические области G_1), для которых

$$P\{(x_1, \ldots, x_n) \in G_1 \mid (s_1, \ldots, s_M) \in S_H\} = \alpha. \quad (1.155)$$

Если в классе этих правил можно найти такое, которое минимизирует вероятность ошибки второго рода или имеет максимальную мощность

$$1-\beta = P\{(x_1, \ldots, x_n) \in G_1^* \mid (s_1, \ldots, s_M) \in S_K\}, \quad (1.155')$$

то это правило называют равномерно наиболее мощным. Можно сузить класс оптимальных правил, вводя условие *несмещенности* (1.119).

Ограничимся здесь лишь одним примером, когда по критерию Неймана — Пирсона необходимо проверить сложную гипотезу о том, что выборка $x_1, \ldots, x_n$ принадлежит нормальному распределению со средним a_0 и *неизвестной дисперсией* σ^2, против сложной альтернативы, что эта выборка принадлежит нормальному распределению со средним $a \neq a_0$ и неизвестной дисперсией. Рассмотрим случайную величину [см. (2.153)]

$$t = t(x_1, \ldots, x_n; a) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (x_i - a) \times$$

$$\times \left\{\frac{1}{n-1} \sum_{k=1}^{n} \left[x_k - \frac{1}{n} \sum_{i=1}^{n} (x_i - a)\right]^2\right\}^{-1/2}, \quad (1.156)$$

подчиняющуюся при независимых нормальных x_i с параметрами (a, σ^2) закону распределения Стьюдента $S_{n-1}(t)$ с $n-1$ степенями свободы [см. (2.154)], который не зависит от a и σ^2. Таким образом, критические области G_1 n-мерного пространства выборок, к которым отнесены выборки $(x_1, \ldots, x_n)$, удовлетворяющие при $a = a_0$ условию

$$\int_{t_1}^{t_2} S_{n-1}(t)\, dt = \alpha, \quad (1.156')$$

соответствуют указанным выше правилам с постоянным уровнем значимости. Можно показать, что при $a > a_0$ среди всех этих правил равномерно наиболее мощным является такое, при котором критической области соответствуют

значения случайной величины t, превосходящие порог t_α (т. е. в (1.156′) $t_2 = \infty$, $t_1 = t_\alpha$).

Иначе говоря, отвергается гипотеза H о том, что среднее равно a_0 при неизвестной дисперсии, если

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - a_0) > \left\{\frac{1}{n-1}\sum_{k=1}^{n}\left[x_k - \frac{1}{n}\sum_{i=1}^{n}(x_i - a_0)\right]^2\right\}^{1/2}\frac{t_\alpha}{\sqrt{n}}, \tag{1.157}$$

где t_α — процентная точка распределения Стьюдента. Это правило выбора решения — равномерно наиболее мощное относительно сложной альтернативы $a > a_0$.

Сравнивая правило (1.157) с аналогичным правилом (1.128) для проверки простой гипотезы о среднем против сложной альтернативы, когда дисперсия нормального распределения известна, замечаем, что в (1.157) неизвестная дисперсия представлена выражением в фигурных скобках, а процентная точка x_α нормального распределения заменена процентной точкой распределения Стьюдента (для заданного α).

Прежде чем вычислять вероятность ошибки второго рода (или мощность $1-\beta$) при использовании правила (1.157), следует иметь в виду, что случайная величина $t(x_1, \ldots, x_n; a_0)$, если x_i, $i = 1, \ldots, n$, нормальны с параметром $a \neq a_0$, имеет *нецентральное* распределение Стьюдента

$$w_1(t; \delta) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n-1}{2}\right)\sqrt{\pi(n-1)}} \times$$

$$\times \int_0^\infty y^{\frac{n}{2}-1}\exp\left\{-\frac{1}{2}\left[y + \left(t\sqrt{\frac{y}{n-1}} - \delta\right)^2\right]\right\} dy, \tag{1.158}$$

где δ — параметр нецентральности, равный

$$\delta = \frac{a - a_0}{\sigma}\sqrt{n}. \tag{1.158′}$$

Используя (1.158), можно записать выражение функции мощности правила (1.157) в виде

$$1 - \beta(\delta) = \int_{t_\alpha}^{\infty} w_1(t; \delta)\, dt, \tag{1.159}$$

где t_α — процентная точка нецентрального распределения Стьюдента.

Если $a < a_0$, то равномерно наиболее мощное правило проверки сложной гипотезы о том, что среднее равно a_0 при неизвестной дисперсии, определяет критическую область неравенством [см. (1.129')]

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - a_0) <$$

$$< -\left\{\frac{1}{n-1}\sum_{k=1}^{n}\left[x_k - \frac{1}{n}\sum_{i=1}^{n}(x_i - a)\right]^2\right\}^{1/2}\frac{t_\alpha}{\sqrt{n}}. \quad (1.160)$$

В том случае, когда альтернатива содержит все действительные значения a, равномерно наиболее мощного правила не существует. Но аналогично (1.131) правило, согласно которому гипотеза H отвергается, если

$$\left|\frac{1}{n}\sum_{i=1}^{n}(x_i - a_0)\right| >$$

$$\geqslant \left\{\frac{1}{n-1}\sum_{k=1}^{n}\left[x_k - \frac{1}{n}\sum_{i=1}^{n}(x_i - a)\right]^2\right\}^{1/2}\frac{t_{\alpha/2}}{\sqrt{n}} \quad (1.161)$$

является наиболее мощным *несмещенным* правилом с заданной вероятностью α ошибки первого рода.

Вероятность ошибки второго рода при этом равна

$$\beta(\delta) = \int_{-t_{\alpha/2}}^{t_{\alpha/2}} w_1(t;\ \delta)\,dt. \quad (1.162)$$

Наконец, в заключение отметим, что обобщение последовательного анализа на случай многих параметров базируется на соображениях, аналогичных тем, которые изложены в § 1.4.6 с усложнениями, связанными с переходом от одномерного к многомерному пространству параметров.

1.5.2. Выборка из многомерного распределения. До сих пор рассматривались задачи проверки гипотез об одном параметре или совокупности параметров *одномерной* функции распределения случайной величины ξ по выборке, элементы которой извлечены из множества возможных значе-

ний этой случайной величины. Более общей является задача проверки гипотез о параметрах *многомерного* распределения совокупности случайных величин $\xi_1, \ldots, \xi_M$. Соответственно элементами многомерной выборки размером n являются группы из N чисел $x_{1k}, x_{2k}, \ldots, x_{Nk}$, $k = 1, \ldots, n$, которые можно рассматривать как компоненты вектора $\mathbf{x}_k$. Тогда выборка является прямоугольной матрицей $\mathbf{X}$ размером $N \times n$:

$$\mathbf{X} = \| x_{ik} \|, \quad i = 1, \ldots, N; \qquad k = 1, \ldots, n. \quad (1.163)$$

Проиллюстрируем обобщение теории проверки гипотез о параметрах многомерного распределения на примере N-мерного нормального распределения. Используя векторную форму записи [см. (2.57) в первой книге], представим многомерную плотность вероятности N зависимых нормальных случайных величин в виде

$$w_N(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{N}{2}} \sqrt{\det M}} \exp\left[-\frac{1}{2} (\mathbf{x} - \mathbf{a})' \mathbf{M}^{-1} (\mathbf{x} - \mathbf{a}) \right], \quad (1.164)$$

где $\mathbf{a}$ — вектор-столбец средних значений; $\mathbf{M}$ — корреляционная матрица, штрих указывает на операцию транспонирования.

Простейшая задача проверки гипотез в этом случае формулируется следующим образом. Имеется одно выборочное значение $\mathbf{x}$ (векторное) из распределения (1.164). Выдвигается простая гипотеза H_0, что это значение принадлежит нормальному распределению с вектором средних $\mathbf{a}_0$ и корреляционной матрицей $\mathbf{M}$, против простой альтернативы H_1, что это значение принадлежит нормальному распределению с вектором средних $\mathbf{a}_1$ и той же самой корреляционной матрицей $\mathbf{M}$.

Логарифм отношения правдоподобия в этом случае имеет следующий вид:

$$\ln l(\mathbf{x}) = \ln \frac{w_N(\mathbf{x} \mid \mathbf{a}_1)}{w_N(\mathbf{x} \mid \mathbf{a}_0)} =$$

$$= -\frac{1}{2} [(\mathbf{x} - \mathbf{a}_1)' \mathbf{M}^{-1} (\mathbf{x} - \mathbf{a}_1) - (\mathbf{x} - \mathbf{a}_0)' \mathbf{M}^{-1} (\mathbf{x} - \mathbf{a}_0)] =$$

$$= \mathbf{x}' \mathbf{M}^{-1} (\mathbf{a}_1 - \mathbf{a}_0) - \frac{1}{2} (\mathbf{a}_1 + \mathbf{a}_0)' \mathbf{M}^{-1} (\mathbf{a}_1 - \mathbf{a}_0). \quad (1.165)$$

Структура этого выражения аналогична структуре (1.66). Обе указанные формулы совпадают при $N = n = 1$.

Правило выбора решения формулируется теперь так: принимается решение γ_1 (среднее равно a_1), если для наблюдаемого вектора $\mathbf{x}$

$$\mathbf{x}'\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0) \geqslant \frac{1}{2}(\mathbf{a}_1+\mathbf{a}_0)'\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0)+\ln c, \quad (1.166)$$

и принимается решение γ_0 (среднее равно $\mathbf{a}_0$), если выполняется неравенство, противоположное (1.166). Как и раньше, величина c определяется в зависимости от выбранного критерия (по располагаемым априорным данным о принадлежности к одному из двух распределений и о функции потерь).

Из (1.166) видно, что процедура проверки гипотезы о среднем значении многомерного нормального распределения сводится к вычислению *линейной функции* компонент вектора результатов наблюдений и сравнению ее с порогом

$$K=\frac{1}{2}(\mathbf{a}_1+\mathbf{a}_0)'\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0)+\ln c. \quad (1.167)$$

Для вычисления условных вероятностей ошибок первого и второго рода можно воспользоваться формулами (1.25) и (1.26), т. е. отображением N-мерного пространства вектора $\mathbf{x}$ на одномерную область отношения правдоподобия. Из (1.165) следует, что $\ln l(\mathbf{x})$ представляет нормальную случайную величину. Поэтому достаточно определить ее среднее и дисперсию при двух гипотезах H_1 и H_0. Имеем

$$m_1\{\ln l(\mathbf{x})\,|\,H_1\}=$$
$$=m_1\{\mathbf{x}'\}\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0)-\frac{1}{2}(\mathbf{a}_1+\mathbf{a}_0)'\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0)=$$
$$=\mathbf{a}_1'\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0)-\frac{1}{2}(\mathbf{a}_1+\mathbf{a}_0)'\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0)=$$
$$=\frac{1}{2}(\mathbf{a}_1-\mathbf{a}_0)'\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0)$$

и, обозначая

$$d^2=(\mathbf{a}_1-\mathbf{a}_0)'\mathbf{M}^{-1}(\mathbf{a}_1-\mathbf{a}_0), \quad (1.168)$$

получим

$$m_1\{\ln l(\mathbf{x})\,|\,H_1\}=\frac{d^2}{2}. \quad (1.169)$$

Аналогичным образом находим

$$m_1\{\ln l(\mathbf{x})\,|\,H_0\}=-\frac{d^2}{2}, \quad (1.170)$$

$$M_2\{\ln l(\mathbf{x})\,|\,H_0\}=M_2\{\ln l(\mathbf{x})\,|\,H_1\}=d^2. \quad (1.171)$$

Тогда вероятности ошибок первого и второго рода равны [см. (1.73) и (1.73′)]

$$\alpha = P\{\ln l(\mathbf{x}) \geqslant \ln c \mid H_0\} = 1 - F\left(\frac{d}{2} + \frac{\ln c}{d}\right), \quad (1.172)$$

$$\beta = P\{\ln l(\mathbf{x}) < \ln c \mid H_1\} = F\left(-\frac{d}{2} + \frac{\ln c}{d}\right). \quad (1.172')$$

Если результаты наблюдений представлены не одним вектором $\mathbf{x}$, а n векторами $\mathbf{x}_k$, $k = 1, \ldots, n$, т. е. матрицей $\mathbf{X}$ размером $N \times n$ [см. (1.163)], то задача проверки гипотез о среднем значении многомерного нормального распределения сводится к рассмотренной выше, если $\mathbf{x}$ заменить на среднее арифметическое $\frac{1}{n}\sum_{k=1}^{n}\mathbf{x}_k$, а матрицу $\mathbf{M}$ — на $\mathbf{M}/n$. Величина d в (1.172) и (1.172′) при этом должна быть заменена на $d_n = d\sqrt{n}$.

1.5.3. Непараметрические методы проверки гипотез о симметрии распределения. В непараметрических случаях, когда неизвестен даже вид распределения, которому принадлежит выборка, правило выбора решения может иногда базироваться на использовании *упорядоченности* выборочных значений. Пусть, например, проверяется гипотеза H о симметрии относительно нуля (четности) функции распределения $w_1(x)$, которому принадлежит выборка $x_1, \ldots, x_n$.

Введем счетчик знаков:

$$u(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leqslant 0. \end{cases} \quad (1.173)$$

Одним из возможных правил выбора решений является следующее: отвергается гипотеза H, т. е. принимается решение (γ_1), что

$$w_1(x) \neq w_1(-x) \quad (1.174)$$

или

$$F_1(x) \neq 1 - F_1(-x), \quad (1.174')$$

если для выборки $x_1, \ldots, x_n$ заданного размера n

$$\sum_{i=1}^{n} u(x_i) > c, \quad (1.175)$$

и принимается решение (γ_0), что распределение $w_1(x)$ симметрично относительно нуля, если выполняется неравенство, обратное (1.175).

Так как элементы выборки x_i независимы и распределены одинаково, то сумма $\sum_{i=1}^{n} u(x_i)$ подчиняется биномиальному закону распределения (см. § 1.2 в первой книге) с параметрами n и $p = P\{x_i > 0\} = 1 - F_1(0)$. Гипотеза H симметрии $w_1(x)$ относительно нуля равносильна утверждению

$$(H)\ F_1(0) = \frac{1}{2}, \tag{1.176}$$

а альтернатива K состоит тогда в том, что

$$(K)\ F_1(0) \neq \frac{1}{2}. \tag{1.176'}$$

Таким образом, сумма в левой части неравенства (1.175) подчиняется биномиальному закону с параметрами $\left(n, \frac{1}{2}\right)$, если справедлива гипотеза H, и этому же закону с параметрами $\left(n, \ p \neq \frac{1}{2}\right)$, если справедлива альтернатива K. Нетрудно записать выражения условных вероятностей ошибок при использовании правила (1.175).

Условная вероятность ошибки первого рода равна [см. (1.22) в первой книге]

$$\alpha = P\{\gamma_1 \mid H\} = P\left\{\sum_{i=1}^{n} u(x_i) > c \mid H\right\} =$$

$$= \sum_{k=[c]+1}^{n} \binom{n}{k} \left(\frac{1}{2}\right)^n = 1 - I_{\frac{1}{2}}(n - [c], \ [c] + 1), \tag{1.177}$$

где $I_q(n - m, \ m + 1)$ — отношение неполной бета-функции к полной бета-функции и $[c]$ — целая часть числа c.

Условная вероятность ошибки второго рода равна

$$\beta = P\{\gamma_0 \mid K\} = P\left\{\sum_{i=1}^{n} u(x_i) \leqslant c \mid K\right\} =$$

$$= \sum_{l=0}^{[c]} \binom{n}{l} p^l (1 - p)^{n-l} = I_{1-p}(n - [c], \ [c] + 1). \tag{1.178}$$

Из (1.177) и (1.178) следует, что при $p = 1 - F_1(0) > 0{,}5$ правило (1.165) *несмещенное*. Действительно, при $q =$

$= 1 - p < \frac{1}{2}$ из неравенства (см. (1.22) в первой книге)

$$B_{1/2}(n-m,\ m+1) = \int\limits_0^{1/2} z^{n-m-1}(1-z)^m\,dz >$$

$$> \int\limits_0^{q} z^{n-m-1}(1-z)^m\,dz = B_q(n-m,\ m+1)$$

следует неравенство

$$I_{1/2}(n-[c]),\ [c]+1) > I_q(n-[c],\ [c]+1),$$

и, следовательно,

$$1-\beta > \alpha.$$

При большом n биномиальное распределение аппроксимируется нормальным (см. § 1.2.2 в первой книге) и формулы (1.177) и (1.178) могут быть переписаны в виде

$$\alpha \sim 1 - F\left(\frac{[c]+1-\frac{n}{2}}{\sqrt{n}/2}\right), \qquad (1.179)$$

$$\beta \sim F\left(\frac{[c]+1-np}{\sqrt{np(1-p)}}\right), \qquad (1.180)$$

где $F(x)$ — интеграл Лапласа.

При заданном уровне значимости α (вероятности ошибки первого рода) порог c в правиле (1.175) определяется из соотношения

$$[c] = \frac{2x_\alpha}{\sqrt{n}} + \frac{n}{2} - 1, \qquad (1.181)$$

где x_α — процентная точка биномиального (или асимптотически нормального при $n \gg 1$) распределения. Подставляя (1.181) в (1.180), получаем

$$\beta \sim F\left[\frac{\frac{2x_\alpha}{n} - \sqrt{n}\left(p-\frac{1}{2}\right)}{\sqrt{p(1-p)}}\right]. \qquad (1.182)$$

Из (1.182) видно, что для несмещенного правила $\left(p > \frac{1}{2}\right)$ при $n \to \infty$ вероятность ошибки второго рода $\beta \to 0$.

Если $p = 1 - F_1(0) < \frac{1}{2}$, то *несмещенным* будет правило, согласно которому гипотеза H отвергается при выполнении неравенства, обратного (1.175).

Тогда

$$\alpha = P\left\{\sum_{i=1}^{n} u(x_i) \leqslant c \mid H\right\} = I_{\frac{1}{2}}(n-[c],\ [c]+1), \quad (1.183)$$

$$\beta = P\left\{\sum_{i=1}^{n} u(x_i) > c \mid K\right\} = 1 - I_{1-p}(n-[c],\ [c]+1) \quad (1.183')$$

и при $p < \frac{1}{2}$ из (1.183) и (1.183′) следует $1-\beta > \alpha$. Для заданного α порог c определяется из равенства

$$[c] = \frac{2x_{1-\alpha}}{\sqrt{n}} + \frac{n}{2} - 1. \quad (1.184)$$

Может быть сформулировано также *двухстороннее* правило выбора решения, согласно которому гипотеза H о симметрии распределения отвергается, если и общее число положительных элементов выборки, и общее число отрицательных элементов превосходит некоторый порог c, т. е.

$$\sum_{i=1}^{n} u(x_i) > c, \quad n - \sum_{i=1}^{n} u(x_i) > c$$

или

$$c < \sum_{i=1}^{n} u(x_i) < n - c. \quad (1.185)$$

Тогда вероятности ошибок равны

$$\alpha = P\left\{c < \sum_{i=1}^{n} u(x_i) < n - c \mid H\right\} =$$

$$= \sum_{k=[c]+1}^{n-[c]-1} \binom{n}{k} \left(\frac{1}{2}\right)^n =$$

$$= I_{\frac{1}{2}}([c]+1,\ n-[c]) - I_{\frac{1}{2}}(n-[c],\ [c]+1), \quad (1.186)$$

$$\beta = 1 - P\left\{c < \sum_{i=1}^{n} u(x_i) < n - c \mid k\right\} = 1 -$$

$$- \sum_{l=[c]+1}^{n-[c]-1} \binom{n}{l} p^l (1-p)^{n-l} =$$

$$= 1 - I_{1-p}([c]+1,\ n-[c]) + I_{1-p}(n-[c],\ [c]+1). \quad (1.187)$$

При $n \gg 1$ правые части этих формул переходят в интегралы Лапласа

$$\alpha \sim 1 - 2F\left(\frac{[c]+1-\frac{n}{2}}{\sqrt{n}/2}\right), \tag{1.188}$$

$$\beta \sim F\left(\frac{[c]+1-np}{\sqrt{np(1-p)}}\right) + F\left(\frac{[c]+1-n(1-p)}{\sqrt{np(1-p)}}\right). \tag{1.189}$$

Фиксируя*) уровень значимости α, находим из (1.188) величину порога c в (1.175):

$$c = \frac{2}{\sqrt{n}} x_{\frac{1+\alpha}{2}} + \frac{n}{2} - 1, \tag{1.190}$$

где $x_{\frac{1+\alpha}{2}}$ — процентная точка нормального распределения.

Подставляя (1.190) в (1.189), получаем

$$\beta \sim F\left[\frac{\frac{2}{\sqrt{n}} x_{\frac{1+\alpha}{2}} + \frac{n}{2}(1-2p)}{\sqrt{np(1-p)}}\right] + F\left[\frac{\frac{2}{\sqrt{n}} x_{\frac{1+\alpha}{2}} - \frac{n}{2}(1-2p)}{\sqrt{np(1-p)}}\right]. \tag{1.191}$$

Вероятности ошибок могут быть уменьшены, если вместо (1.175) использовать более мощное правило: отвергается гипотеза о симметрии распределения, если для выборки $x_1, \ldots, x_n$ заданного размера n

$$\sum_{j=1}^{n} \sum_{i=j}^{n} u(x_i + x_j) > c, \tag{1.192}$$

где $u(x)$ — функция, определенная согласно (1.173) (см. [7, § 6.9]).

Задачи

1.1. Пусть H_0 — простая гипотеза, состоящая в том, что выборка $x_1, \ldots, x_n$ получена из экспоненциального распределения

$$w_1(x) = e^{-x}, \quad x > 0, \tag{1}$$

---

*) Так как $n \gg 1$, то $[c] \approx c$, и не имеет существенного значения, что правая часть (1.190) не при любом α будет целым числом.

а H_1 — простая альтернатива, состоящая в том, что эта выборка получена из одностороннего нормального распределения

$$w_1(x) = \sqrt{\frac{2}{\pi}}\, e^{-\frac{x^2}{2}}, \quad x > 0. \tag{2}$$

Доказать следующее оптимальное правило выбора решения: отвергается гипотеза H_0, если

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - 1)^2 < 1 + \ln\frac{2}{\pi} - \frac{2\ln c}{n}, \tag{3}$$

где c — константа, характеризующая выбранный критерий качества. Показать, что при $n \to \infty$ имеют место следующие асимптотические формулы для условных вероятностей ошибок первого и второго рода:

$$\alpha \sim 1 - F\left(\frac{\ln c + \frac{n}{2}\ln\frac{\pi}{2}}{\sqrt{2n}}\right), \tag{4}$$

$$\beta \sim F\left[\frac{\ln c - \frac{n}{2}\left(2\sqrt{\frac{2}{\pi}} + \ln\frac{2}{\pi} - 1\right)}{\sqrt{\left(\frac{3}{2} - \frac{2}{\pi} - \sqrt{\frac{2}{\pi}}\right)n}}\right]. \tag{4'}$$

1.2. Рассматриваются последовательности N независимых испытаний (см. § 1.2 в первой книге). Пусть x_i — число появления событий в i-й последовательности ($i = 1, \ldots, n$). Выдвигается простая гипотеза H_0 о том, что априорная вероятность появления события в каждом испытании равна p_0, против простой альтернативы H_1, что эта вероятность равна p_1 $(p_1 > p_0)$. Доказать следующее оптимальное правило выбора решения: отвергается гипотеза H_0, если

$$\frac{1}{n}\sum_{i=1}^{n} x_i > \frac{\ln\left[\left(\frac{1-p_0}{1-p_1}\right)^N c^{\frac{1}{n}}\right]}{\ln\frac{p_1(1-p_0)}{p_0(1-p_1)}} = K, \tag{5}$$

где c — константа, характеризующая выбранный критерий. Показать, что вероятности ошибок первого и второго рода для правила (5) равны соответственно

$$\alpha = 1 - \frac{B_{1-p_0}(Nn - [nK],\ [nK]+1)}{B(Nn - [nK],\ [nK]+1)}, \tag{6}$$

$$\beta = \frac{B_{1-p_1}(Nn - [nK],\ [nK]+1)}{B(Nn - [nK],\ [nK]+1)}, \tag{6'}$$

где $B_q(x, y)$ и $B(x, y)$ — неполная и полная бета-функции.

1.3. Пусть H_0 — простая гипотеза, состоящая в том, что выборка $x_1, \ldots, x_n$ получена из релеевского распределения

$$w_1(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \tag{7}$$

с параметром $\sigma = \sigma_0$, а H_1 — простая альтернатива, состоящая в том, что эта выборка получена из того же распределения с параметром $\sigma = \sigma_1 > \sigma_0$. Доказать следующее оптимальное правило выбора решения: отвергается гипотеза H_0, если

$$\frac{1}{n} \sum_{i=1}^{n} x_i^2 > \frac{2\sigma_0^2\sigma_1^2}{\sigma_1^2 - \sigma_0^2} \ln\left(\frac{\sigma_1^2}{\sigma_0^2} c^{\frac{1}{n}}\right) = K, \tag{8}$$

где c — константа, характеризующая выбранный критерий качества. Показать, что вероятности ошибок первого и второго рода для правила (8) равны соответственно

$$\alpha = \frac{\Gamma\left(n, \frac{nK}{\sigma_0^2}\right)}{\Gamma(n)}, \tag{9}$$

$$\beta = 1 - \frac{\Gamma\left(n, \frac{nK}{\sigma_1^2}\right)}{\Gamma(n)}, \tag{9'}$$

где $\Gamma(n, x)$ — неполная гамма-функция. Объяснить совпадение формул (9) и (9′) с (1.98) и (1.98′) при $\sigma^2 = \frac{1}{2\lambda}$.

1.4. Показать, что дисперсии логарифма одномерного отношения правдоподобия для двух нормальных распределений со средними a_1 и a_0 и одинаковыми дисперсиями σ^2 равны [см. (1.47) и (1.48)]:

$$M_{20} = M_{21} = \frac{(a_1 - a_0)^2}{\sigma^2}. \tag{10}$$

1.5. Пусть H_0 — простая гипотеза, состоящая в том, что x получено из нормального распределения

$$w_1(x \mid H_0) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \tag{11}$$

а H_1 — простая альтернатива, состоящая в том, что эта величина определена из бимодального распределения

$$w_1(x \mid H_1) = \frac{1}{2\sqrt{\pi}} [e^{-(x-a)^2} + e^{-(x+a)^2}]. \tag{12}$$

Доказать следующее оптимальное правило выбора решения: отвергается гипотеза H_0, если

$$l(x) = \frac{1}{\sqrt{2}} e^{a^2} [e^{-\frac{(x-2a)^2}{2}} + e^{-\frac{(x+2a)^2}{2}}] > c. \tag{13}$$

Обратить внимание на то, что область действительной оси x, определяемая неравенством (13), двухсвязная.

1.6. Пусть $\mathbf{x}$ — выборочное значение (векторное) из многомерного нормального распределения с корреляционной матрицей $\sigma^2\mathbf{I}$ (где $\mathbf{I}$ — единичная матрица). Выдвигается гипотеза H_0, что среднее (векторное) равно $\mathbf{a}_0$, против альтернативы H_1, что оно равно $\mathbf{a}_1$. Показать, что оптимальное по критерию максимального правдоподобия разбиение пространства выборок производится гиперплоскостью, перпендикулярной линии, соединяющей точки $\mathbf{x}=\mathbf{a}_0$ и $\mathbf{x}=\mathbf{a}_1$, и делящей эту линию пополам.

1.7. Пусть $\mathbf{x}$ — выборочное значение (векторное) из многомерного нормального распределения. Выдвигается гипотеза H_0, что это значение принадлежит нормальному распределению с вектором средних $\mathbf{a}$ и корреляционной матрицей $\mathbf{M}_0$, против альтернативы H_1, что это значение принадлежит нормальному распределению с тем же самым вектором средних и корреляционной матрицей $\mathbf{M}_1$. Показать, что оптимальное правило выбора решения формулируется следующим образом: принимается решение γ_1 (корреляционная матрица равна $\mathbf{M}_1$), если для наблюдаемого вектора $\mathbf{x}$

$$(\mathbf{x}-\mathbf{a})'(\mathbf{M}_1-\mathbf{M}_0)(\mathbf{x}-\mathbf{a}) \geqslant 2\ln c + \ln\frac{\det \mathbf{M}_1}{\det \mathbf{M}_0}, \tag{14}$$

и принимается решение γ_0 (корреляционная матрица равна $\mathbf{M}_0$), если выполняется неравенство, обратное (14).

Показать, что, совершая замену переменных

$$\mathbf{y}=\mathbf{f}'(\mathbf{x}-\mathbf{a}), \tag{15}$$

где матрица $\mathbf{f}$ определяется из

$$\mathbf{M}_1\mathbf{f}=\mathbf{M}_0\mathbf{f}\mathbf{\Lambda} \tag{16}$$

и $\mathbf{\Lambda}$ — диагональная матрица, элементы $\lambda_1, \ldots, \lambda_N$ которой являются корнями уравнения

$$\det[\mathbf{M}_1-\lambda\mathbf{M}_0]=0, \tag{17}$$

можно неравенство (14) привести к виду

$$\sum_{i=1}^{N}\left(1-\frac{1}{\lambda_i}\right)y_i^2 \geqslant 2\ln c+\sum_{i=1}^{N}\ln\lambda_i. \tag{18}$$

1.8. Проверяется простая гипотеза H, что дисперсия нормальной случайной величины с нулевым средним равна σ_0^2, против сложной альтернативы, что дисперсия $\sigma^2 \neq \sigma_0^2$. Доказать, что правило, согласно которому отвергается гипотеза H, если

$$\sum_{k=1}^{n} x_k^2 \geqslant \sigma_0^2\chi_\alpha^2, \tag{19}$$

где χ_α^2 — процентная точка χ^2-распределения с n степенями свободы, является равномерно наиболее мощным относительно сложной альтернативы, если $\sigma^2>\sigma_0^2$.

1.9. Проверяется простая гипотеза H, что параметр λ экспоненциального распределения равен λ_0, против сложной альтернативы, что $\lambda \neq \lambda_0$. Доказать, что правило, согласно которому отвергается гипотеза H, если

$$\sum_{k=1}^{n} x_k \leqslant \frac{1}{2\lambda_0} \chi^2_{1-\alpha}, \tag{20}$$

где $\chi^2_{1-\alpha}$ — процентная точка χ^2-распределения с $2n$ степенями свободы, является равномерно наиболее мощным относительно сложной альтернативы, если $\lambda > \lambda_0$.

1.10. Доказать, что условный максимум функционала

$$\Phi[\varphi(x)] = \int_{-\infty}^{\infty} w(x) \ln \varphi(x)\, dx \tag{21}$$

при условии $\varphi(x) > 0$ и $\int_{-\infty}^{\infty} \varphi(x)\, dx = 1$ достигается при

$$\varphi(x) = w(x), \tag{22}$$

т. е. что

$$\int_{-\infty}^{\infty} w(x) \ln w(x)\, dx \geqslant \int_{-\infty}^{\infty} w(x) \ln \varphi(x)\, dx. \tag{23}$$

ЛИТЕРАТУРА

1. Башаринов А. Е., Флейшман Б. С. Методы статистического последовательного анализа и их приложения. Изд-во «Советское радио», 1962, гл. 1, 3.
2. Большаков И. А., Гуткин Л. С., Левин Б. Р., Стратонович Р. Л. Математические основы современной радиоэлектротехники. Изд-во «Советское радио», 1968, гл. 6.
3. Вальд А. Последовательный анализ. Физматгиз, 1960, гл. 1—4.
4. Давенпорт В. Б., Рут В. Л. Введение в теорию случайных сигналов и шума. Пер. с англ., под ред. Р. Л. Добрушина. Изд-во иностранной литературы, 1960, § 14.2.
5. Дунин-Барковский И. В., Смирнов Н. В. Курс теории вероятностей и математической статистики. Изд-во «Наука», 1965, гл. VII.
6. Крамер Г. Математические методы статистики. Пер. с англ., под ред. А. Н. Колмогорова. Изд-во иностранной литературы, 1948, гл. 25, 26, 35.
7. Леман Э. Проверка статистических гипотез. Пер. с англ. Ю. В. Прохорова. Изд-во «Наука», 1964, гл. 1, 3.

8. Линник Ю. В. Статистические задачи с мешающими параметрами. Изд-во «Наука», 1966, гл. III, IV.
9. Миддлтон Д. Введение в статистическую теорию связи. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1962, гл. 18.
10. Миддлтон Д. Очерки теории связи. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1966, гл. 1, 2.
11. Хазен Э. М., Бененсон З. М. Последовательные решающие правила в некоторых задачах о различении многих гипотез и сложных гипотез. Тезисы сообщений на Международном конгрессе математиков, секция 11, М., 1966.
12. Хельстром К. Статистическая теория обнаружения сигналов. Пер. с англ., под ред. Ю. Б. Кобзарева. Изд-во иностранной литературы, 1963, гл. 3.
13. Шэннон К. Математическая теория связи. В сборнике переводов. «Работы по теории информации и кибернетике», под ред. Р. Л. Добрушина и О. Б. Лупанова. Изд-во иностранной литературы, 1963.
14. Wald A. Statistical Decision Functions, New York, John Wiley Inc., 1950, гл. 1, 3.

ГЛАВА ВТОРАЯ

СТАТИСТИКА СЛУЧАЙНЫХ ВЕЛИЧИН

2.1. ВЫБОРКА И ЕЕ ХАРАКТЕРИСТИКИ

2.1.1. Эмпирическая функция распределения. Вернемся к § 1.1.4 и рассмотрим снова некоторый случайный эксперимент и случайную величину, связанную с ним в том смысле, что возможные значения этой случайной величины представляют результаты отдельных наблюдений. Последовательность n наблюдений характеризуется *выборкой* размера n, элементами которой являются указанные возможные значения случайной величины ξ.

Будем рассматривать выборки, которые получаются в результате независимых наблюдений. Процесс извлечения таких выборок называют *простым случайным выбором*.

Если известны интегральная функция распределения $F_1(x)$ или плотность вероятности $w_1(x)$ случайной величины ξ, то говорят, что выборка $x_1, \ldots, x_n$ принадлежит распределению $F_1(x)$ или $w_1(x)$. Перегруппировав элементы выборки путем расстановки их в возрастающем порядке так, что $x_i \geqslant x_j$ при $i > j$, получают *упорядоченную выборку*, которую называют *статистическим* (или вариационным) рядом. Отдельные члены статистического ряда называют иногда порядковыми *статистиками*. Примерами порядковых статистик является наименьшее значение x_1, наибольшее x_n, выборочная медиана $x_{\frac{n}{2}+1}$ (если n — четное).

Пусть $\nu(x)$ — число выборочных значений, не превосходящих некоторого порога x. Ступенчатая функция

$$F_1^*(x) = \frac{\nu(x)}{n} = \frac{1}{n} \sum_k u(x - x_k), \tag{2.1}$$

где $u(x)$ — единичный скачок [см. (1.173)], представляет частоту события $\xi \leqslant x$ в последовательности n наблюдений. Ее называют *эмпирической функцией распределения выборки*.

Эта функция представляет статистический аналог функции распределения случайной величины, но не совпадает с последней (рис. 12). Эмпирическая функция распределения $F_1^*(x)$ при $n \to \infty$ *сходится по вероятности* к гипотетической функции распределения $F_1(x)$, если выборка, по которой построена эмпирическая функция, извлечена из распределения $F_1(x)$ (теорема Гливенко, см. [3]), т. е. при любом x и $\varepsilon > 0$

$$\lim_{n\to\infty} P\{|F_1(x) - F_1^*(x)| \geqslant \varepsilon\} = 0. \qquad (2.2)$$

Когда осуществляется выбор из непрерывного распределения и число выборочных значений очень велико, можно

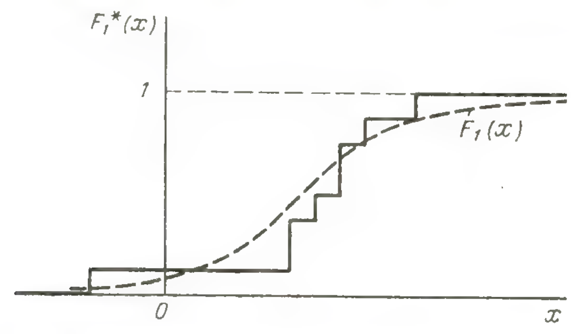


Рис. 12. Эмпирическая функция распределения.

производить *группирование*. При этом область возможных значений случайной величины разделяют на N непересекающихся интервалов и объединяют выборочные значения, попадающие в один и тот же интервал. На каждом из этих интервалов как на основании строят прямоугольник высотой $\frac{\nu_i}{nh_i}$, где h_i — длина интервала, ν_i — число выборочных значений, попавших в этот i-й интервал; n — размер выборки. Полученную таким образом ступенчатую функцию

$$w_1^*(x) = \frac{1}{n}\sum_{i=1}^{N}\frac{\nu_i(x)}{h_i}, \qquad (2.3)$$

$$\nu_i(x) = \begin{cases} \nu_i, & x_i \leqslant x < x_{i+1}, \\ 0, & x < x_i,\ x \geqslant x_{i+1}, \end{cases} \qquad (2.3')$$

$$h_i = x_{i+1} - x_i \qquad (2.3'')$$

называют *гистограммой выборки*. Она служит статистическим аналогом плотности вероятности случайной величины (рис. 13).

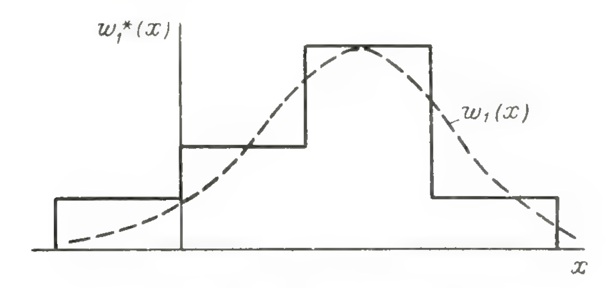


Рис. 13. Гистограмма.

2.1.2. Выборочные моменты. Эмпирическая функция распределения содержит *всю* статистическую информацию, которая накоплена в процессе извлечения выборки. Как известно из теории вероятностей, случайную величину можно характеризовать не только функцией распределения, но и числами — моментами распределения. Аналогично статистические свойства выборки можно характеризовать не только эмпирической функцией распределения, но и более грубо — несколькими числами — *выборочными моментами*. Выборочный момент k-го порядка определяется по формуле

$$m_k^* = \frac{1}{n} \sum_{i=1}^{n} x_i^k, \tag{2.4}$$

т. е. он равен среднему арифметическому k-х степеней выборочных значений. В отличие от моментов случайной величины, выборочные моменты обозначают звездочкой.

Следует всегда отличать выборочные моменты от моментов случайной величины. Последние определяются через априорную функцию распределения $w_1(x)$ по известной формуле (см. (2.77) в первой книге)

$$m_k = \int_{-\infty}^{\infty} x^k w_1(x)\, dx. \tag{2.5}$$

Для дискретной случайной величины, принимающей значения x_i, с вероятностью p_i

$$m_k = \sum_i x_i^k p_i. \tag{2.5'}$$

Таким образом, моменты случайной величины вычисляются по известным априорным характеристикам, в то время как выборочные моменты — на основании наблюденных значений.

Рассмотрим выборочные моменты первых четырех порядков. В соответствии с общим определением выборочный момент первого порядка, или просто *выборочное среднее*, равен среднему арифметическому выборочных данных, т. е.

$$m_1^* = \frac{1}{n} \sum_{i=1}^{n} x_i. \tag{2.6}$$

Выборочное среднее характеризует расположение выборки на действительной оси, указывает область, в которой группируются выборочные значения. Разность $x_i - m_1^*$ называют *отклонением* выборочного значения от выборочного среднего. Выборочные моменты отклонения называются центральными и обозначаются символами M_k^*, $k = 2, 3, \ldots$ В соответствии с (2.4)

$$M_k^* = \frac{1}{n} \sum_{i=1}^{n} (x_i - m_1^*)^k. \tag{2.7}$$

Выборочная дисперсия M_2^* (центральный выборочный момент второго порядка) является мерой рассеяния выборочных значений относительно выборочного среднего:

$$M_2^* = \frac{1}{n} \sum_{i=1}^{n} (x_i - m_1^*)^2. \tag{2.8}$$

С выборочными центральными моментами третьего и четвертого порядка связаны выборочный коэффициент асимметрии k^* и выборочный коэффициент эксцесса γ^*:

$$\mathrm{k}^* = \frac{M_3^*}{(M_2^*)^{3/2}}, \quad \gamma^* = \frac{M_4^*}{(M_2^*)^2}. \tag{2.9}$$

2.1.3. Функция правдоподобия. Каждая выборка размером n, т. е. совокупность n чисел, представляет точку в n-мерном пространстве. Если мысленно представить все возможные исходы экспериментов, то совокупности всевозможных выборок заданного размера n соответствует некоторая область n-мерного пространства выборок. Будем рассматривать не индивидуально каждую выборку, а совокупность всех выборок с фиксированным размером n, принадлежащих априорному распределению $w_1(x)$. Плотность вероятности выборки с независимыми элементами равна [см. (1.6)]

$$W_n(x_1, \ldots, x_n) = \prod_{k=1}^{n} w_1(x_k). \tag{2.10}$$

Вероятность появления выборки $x_1, \ldots, x_n$ из дискретного распределения

$$P_n(x_1, \ldots, x_n) = \prod_{k=1}^{n} p(x_k), \tag{2.10'}$$

где $p(x_k)$ — вероятность того, что $\xi = x_k$.

Совместное распределение $W_n(x_1, \ldots, x_n)$ выборочных значений, как уже указывалось в первой главе, называют *функцией правдоподобия выборки*. Зная распределение $w_1(x)$, можно по формуле (2.10) найти функцию правдоподобия и по известным правилам теории вероятностей (см. гл. 3 в первой книге) определить распределение *любой функции* $g(x_1, \ldots, x_n)$ выборочных значений.

Функции выборочных значений часто называют *статистиками*. Примером статистики являются выборочные моменты. Выборочные значения и их функции, в частности выборочные моменты, являются *случайными величинами*, в то время как числовые характеристики (моменты) исходного распределения $w_1(x)$ являются постоянными числами, хотя все или некоторые из них могут быть неизвестными *).

Когда накоплены статистические данные (выборки), можно использовать выборочные значения и их характеристики в форме эмпирических распределений, гистограмм, выборочных моментов и т. п. для описательных целей.

---

*) Из (2.4) следует, что при соответствующих ограничениях на моменты исходного распределения, указываемые центральной предельной теоремой (см. § 3.4 в первой книге), выборочные моменты m_k^* при большом размере выборки *асимптотически нормальны.*

Однако задачи математической статистики не ограничиваются описанием. Главным является получение статистических выводов о неизвестном или неполностью известном распределении случайной величины на основании ее выборочных значений.

2.1.4. Закон больших чисел. Рассмотрим выборки $x_1, \ldots, x_n$ из распределения $w_1(x)$ случайной величины ξ, имеющей конечную дисперсию σ^2 (а следовательно, и конечное среднее a). Составим выборочное среднее

$$m_1^* = \frac{1}{n} \sum_{i=1}^{n} x_i$$

и вычислим его среднее по пространству выборок:

$$m_1\{m_1^*\} = \frac{1}{n} \sum_{i=1}^{n} m_1\{x_i\} = \frac{1}{n} \sum_{i=1}^{n} a = a, \qquad (2.11)$$

так как $m_1\{x_i\} = a$.

Итак, среднее значение от выборочного среднего при любом размере выборки в точности совпадает со средним значением a случайной величины ξ. Это, конечно, не значит, что любое конкретное выборочное среднее будет совпадать с величиной a. Выборочное среднее — величина случайная, но ее среднее в точности совпадает со средним исходного распределения.

Найдем дисперсию выборочного среднего:

$$M_2\{m_1^*\} = \frac{1}{n^2} \sum_{i=1}^{n} M_2\{x_i\} = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}. \qquad (2.11')$$

Здесь использовано то обстоятельство, что дисперсия суммы равна сумме дисперсии, если слагаемые независимы, и что

$$M_2\{x_i\} = \sigma^2.$$

Таким образом, дисперсия выборочного среднего равна дисперсии случайной величины ξ, деленной на размер выборки. По мере того как увеличивается размер выборки, дисперсия выборочного среднего уменьшается и в пределе при $n \to \infty$ стремится к нулю.

Рассмотрим вероятность того, что выборочное среднее отличается от среднего значения исходного распределения

по модулю на величину, бо́льшую, чем произвольное $\varepsilon > 0$, и исследуем, как меняется эта вероятность, когда размер выборки n неограниченно возрастает. В силу неравенства Чебышева (см. (2.92) в первой книге) эта вероятность не должна превосходить дисперсии выборочного среднего, деленной на ε^2, т. е.

$$P\{|m_1^* - a| \geqslant \varepsilon\} \leqslant \frac{\sigma^2}{n\varepsilon^2}. \qquad (2.12)$$

Из (.212) непосредственно следует, что предел рассматриваемой вероятности, когда размер выборки стремится к бесконечности, равен нулю:

$$\lim_{n\to\infty} P\{|m_1^* - a| \geqslant \varepsilon\} = 0. \qquad (2.13)$$

Предел вероятности противоположного неравенства равен единице:

$$\lim_{n\to\infty} P\{|m_1^* - a| < \varepsilon\} = 1. \qquad (2.13')$$

Формулы (2.13) и (2.13′) являются аналитическим выражением *закона больших чисел*: по мере увеличения размера выборки вероятность того, что выборочное среднее будет сколь угодно мало отличаться от среднего значения случайной величины ξ, стремится к единице.

Следовательно, закон больших чисел можно сформулировать так: выборочное среднее сходится по вероятности к среднему значению исходного распределения (см. § 3.5 в первой книге).

Аналогично можно доказать, что любой выборочный момент k-го порядка сходится по вероятности, когда размер выборки неограниченно возрастает, к соответствующему k-му моменту исходного распределения, если только существует момент порядка $2k$ этого распределения.

Из закона больших чисел следует, что выборочные характеристики могут служить *оценками* соответствующих характеристик исходного распределения. Конечно, достаточно хорошее совпадение выборочных характеристик с параметрами исходного распределения получается только при выборках очень большого размера, которые не всегда удается реализовать. Наоборот, для экономии времени и средств желательно ограничиться возможно меньшим размером выборки, но тогда нельзя гарантировать большой точности.

Возникает задача изучения свойств этих оценок при конечном размере выборки.

Закон больших чисел может быть записан в более общей, чем (2.13), форме:

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} f(x_i) - m_1\{f(x_i)\}\right| \geqslant \varepsilon\right\} = 0, \quad (2.13'')$$

где $f(x)$ — заданная функция и $m_2\{f(x_i)\} < \infty$. При $n \to \infty$ распределение $\frac{1}{n}\sum_{i=1}^{n} f(x_i)$ асимптотически нормальное.

Заметим попутно, что соотношение (2.13″) может быть использовано для статистической оценки величины интеграла

$$\int_{-\infty}^{\infty} f(x)\, w(x)\, dx = m_1\{f(x)\} \approx \frac{1}{n}\sum_{k=1}^{n} f(x_k),$$

для чего нужно располагать выборкой достаточно большого размера из распределения $w(x)$. Этот метод приближенного вычисления интегралов известен как вычислительный метод Монте-Карло.

2.2. ОЦЕНКА ПАРАМЕТРОВ ОДНОМЕРНОЙ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

2.2.1. Точечные оценки. Многие практически важные задачи сводятся к получению статистических выводов относительно одномерной функции распределения, вид которой *известен*, но все или несколько параметров этого распределения неизвестны. Один вид статистических выводов связан с определением *точечных оценок* параметров, т. е. с нахождением таких функций выборочных значений (статистик)

$$\hat{\vartheta}_n^{(i)} = g^{(i)}(x_1, \ldots, x_n), \quad (2.14)$$

которые являются наилучшими (в некотором смысле) оценками неизвестных параметров ϑ_i, $i = 1, \ldots, m$, исходной функции распределения $w_1(x;\ \vartheta_1, \ldots, \vartheta_m)$.

Ограничимся сначала случаем, когда указанное распределение содержит один параметр ϑ. Укажем на общую классификацию точечных оценок параметра одномерного распределения по некоторым свойствам этих оценок.

Важнейшими из этих свойств являются: состоятельность, достаточность, несмещенность и эффективность. Рассмотрим каждое из этих свойств в отдельности. Предварительно заметим, что сам оцениваемый параметр ϑ может быть случайной величиной и характеризоваться априорным распределением $w_1(\vartheta)$. Часто, однако, информация об априорном распределении отсутствует или заранее известно, что оцениваемый параметр не является случайным. В соответствии с этими двумя ситуациями определяются:

1) условная оценка параметра ϑ

$$\hat{\vartheta}_n = g(x_1, \ldots, x_n), \tag{2.15}$$

определяемая по выборочным значениям $x_1, \ldots, x_n$ в предположении, что они принадлежат распределению $w_1(x \mid \vartheta)$ с фиксированным *) значением параметра ϑ;

2) безусловная оценка параметра ϑ

$$\hat{\Theta}_n = \int_{-\infty}^{\infty} w_1(\vartheta)\, g(x_1, \ldots, x_n)\, d\vartheta, \tag{2.15'}$$

получаемая усреднением условной оценки по всем возможным значениям ϑ.

Отметим, что и условная, и безусловная оценки как функции выборочных значений $x_1, \ldots, x_n$ представляют случайные величины, распределения которых определяются условной и безусловной функциями правдоподобия выборки:

$$W_n(x_1, \ldots, x_n \mid \vartheta)$$

и

$$W_n(x_1, \ldots, x_n) = \int_{-\infty}^{\infty} w_1(\vartheta)\, W_n(x_1, \ldots, x_n \mid \vartheta)\, d\vartheta. \tag{2.16}$$

2.2.2. Состоятельность. Условная оценка $\hat{\vartheta}_n$ называется состоятельной оценкой фиксированного параметра ϑ, если она сходится по вероятности к оцениваемому параметру при неограниченном увеличении размера выборки n, т. е. при $\varepsilon > 0$

$$\lim_{n\to\infty} P\{|\hat{\vartheta}_n - \vartheta| \geqslant \varepsilon\} = 0. \tag{2.17}$$

---

*) Обозначение $w(x \mid \vartheta)$ и ему аналогичные используются всегда, когда ϑ — случайный параметр и рассматривается условная плотность x при фиксированном ϑ. Однако обозначение $w(x \mid \vartheta)$, наряду с $w(x, \vartheta)$, часто используется нами и в том случае, когда ϑ — неизвестный (неслучайный) параметр.

Аналогично безусловная оценка $\hat{\Theta}_n$ называется состоятельной оценкой случайного параметра ϑ, если она сходится по вероятности к среднему значению $m_1\{\vartheta\}$ оцениваемого параметра при неограниченном увеличении размера выборки, т. е. если при $\varepsilon > 0$

$$\lim_{n\to\infty} P\{|\hat{\Theta}_n - m_1\{\vartheta\}| \geqslant \varepsilon\} = 0. \tag{2.17'}$$

Закон больших чисел устанавливает состоятельность условной оценки специального вида — выборочного среднего, которое сходится по вероятности к среднему значению исходного распределения. Как указывалось, закон больших чисел может быть обобщен: любой выборочный момент сходится при определенных условиях к соответствующему моменту исходного распределения. Поэтому можно утверждать, что при указанных условиях любой выборочный момент является состоятельной оценкой соответствующего момента исходного распределения.

2.2.3. Несмещенность. Условная оценка $\hat{\vartheta}_n = g(x_1, \ldots x_n)$ называется несмещенной оценкой параметра ϑ, если среднее значение этой оценки по совокупности выборок заданного размера в точности равно оцениваемому параметру, т. е. если

$$m_1\{\hat{\vartheta}_n\} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) W_n(x_1, \ldots, x_n | \vartheta)\, dx_1 \ldots dx_n = \vartheta \tag{2.18}$$

при любом n.

Аналогично *) безусловная оценка $\hat{\Theta}_n$ называется несмещенной оценкой случайного параметра ϑ, если при любом n

$$m_1\{\hat{\Theta}_n\} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \hat{\Theta}_n W_n(x_1, \ldots, x_n)\, dx_1 \ldots dx_n = m_1\{\vartheta\}. \tag{2.19}$$

*) Символы m_1 в начале и в конце формулы (2.19) относятся к усреднению по различным совокупностям: оценка $\hat{\Theta}_n$ усредняется по пространству выборок заданного размера, а параметр ϑ — по его возможным значениям. (Это замечание следует иметь в виду и в дальнейшем.)

Разность

$$m_1\{\hat{\vartheta}_n\}-\vartheta=b_n(\vartheta) \qquad (2.20)$$

называют *смещением* условной оценки (или систематической ошибкой). Для безусловной оценки смещение определяется следующим образом:

$$m_1\{\hat{\Theta}_n\}-m_1\{\vartheta\}=B_n(\vartheta). \qquad (2.21)$$

Заметим, что смещение $b_n(\vartheta)$ легко устранимо, если оно является линейной функцией

$$b_n(\vartheta)=a\vartheta+b, \qquad (2.22)$$

где a и b — любые действительные числа (одно из них может быть равным нулю). Если имеет место указанное смещение, то, заменяя оценку $\hat{\vartheta}_n$ на $\frac{\hat{\vartheta}_n-b}{a+1}$, получаем несмещенную оценку.

Примером несмещенной оценки среднего значения a для произвольного распределения является выборочное среднее $\hat{a}_n=\frac{1}{n}\sum_{i=1}^{n}x_i$, так как

$$m_1\{\hat{a}_n\}=\frac{1}{n}\sum_{i=1}^{n}m_1\{x_i\}=a. \qquad (2.23)$$

Таким образом, выборочное среднее — состоятельная и несмещенная оценка среднего значения исходного распределения. Однако из состоятельности не следует несмещенность оценки. Так, при определенных условиях выборочная дисперсия $\hat{\sigma}_n^2=M_2^*$ представляет состоятельную оценку дисперсии σ^2 исходного распределения. Однако

$$m_1\{\hat{\sigma}_n^2\}=\frac{1}{n}\sum_{i=1}^{n}m_1\{(x_i-m_1^*)^2\}=$$

$$=\frac{1}{n}\sum_{i=1}^{n}[m_1\{(x_i-a)^2\}+m_1\{(m_1^*-a)^2\}-$$

$$-2m_1\{(x_i-a)(m_1^*-a)\}]=$$

$$=\frac{1}{n}\sum_{i=1}^{n}\left(\sigma^2+\frac{\sigma^2}{n}-2\frac{\sigma^2}{n}\right)=\sigma^2\left(1-\frac{1}{n}\right), \qquad (2.24)$$

т. е. оценка $\hat{\sigma}_n^2$ — смещенная. Величина смещения равна $-\frac{\sigma^2}{n}$. При небольшом размере выборки смещение может быть существенным. Например, при $n=3$ оно достигает 33%. Так как в рассматриваемом случае смещение — линейная функция, то в соответствии с (2.20) при $a = -\frac{1}{n}$, $b=0$ находим несмещенную оценку дисперсии:

$$\frac{1}{1-\frac{1}{n}} M_2^* = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - m_1^*)^2. \qquad (2.25)$$

Состоятельность оценки не нарушается, так как $\lim\limits_{n\to\infty}\left(1-\frac{1}{n}\right)=1$.

Заметим, что смещенность оценки $\hat{\sigma}_n^2$ является платой за незнание среднего значения a исходного распределения. Если величина a известна, то

$$m_1\{\hat{\sigma}_n^2\} = \frac{1}{n}\sum_{i=1}^{n} m_1\{(x_i-a)^2\} = \sigma^2$$

и указанная оценка является несмещенной.

Иногда рассматривают оценки, для которых равенства (2.18) и (2.19) выполняются не при конечных размерах выборки, а лишь при неограниченном увеличении n. Оценки (условные или безусловные), для которых

$$\lim_{n\to\infty} m_1\{\check{\vartheta}_n\} = \vartheta \qquad (2.26)$$

или

$$\lim_{n\to\infty} m_1\{\hat{\Theta}_n\} = m_1\{\vartheta\}, \qquad (2.27)$$

называют *асимптотически несмещенными*. Оценка (2.24) является примером асимптотически несмещенной оценки, так как смещение $b_n(\sigma^2) = -\frac{\sigma^2}{n}$ стремится к нулю при $n\to\infty$. Конечно, несмещенная оценка всегда асимптотически несмещенная, но не наоборот.

2.2.4. Достаточность. До сих пор предполагалось, что получение оценок связано с предварительным извлечением выборки заданного размера n. Однако в некоторых случаях

для оценки параметра ϑ нет необходимости знать *каждый* из элементов выборки x_i в отдельности, а достаточно иметь одну или несколько функций $g^{(l)}(x_1, \ldots, x_n)$, $l = 1, \ldots, k$ $(k \leqslant n)$, выборочных значений. Эти функции называют *достаточными оценками* (или *достаточными статистиками*) параметра ϑ. Ограничиваясь достаточными статистиками, можно иногда существенно сократить процесс накопления данных, необходимых для оценки неизвестного параметра.

Необходимое и достаточное условие того, чтобы $g(x_1, \ldots, x_n)$ была достаточной статистикой, состоит в возможности факторизации функции правдоподобия выборки, т. е. представления $W_n(x_1, \ldots, x_n \mid \vartheta)$ в виде произведения двух неотрицательных сомножителей:

$$W_n(x_1, \ldots, x_n \mid \vartheta) = f[g(x_1, \ldots, x_n), \vartheta]\, h(x_1, \ldots, x_n), \quad (2.28)$$

первый из которых зависит от ϑ и от $g(x_1, \ldots, x_n)$, а второй — только от выборочных значений $x_1, \ldots, x_n$ (и не зависит от оцениваемого параметра ϑ).

Если $g(x_1, \ldots, x_n)$ — достаточная оценка случайного параметра ϑ, то апостериорное распределение $W_1(\vartheta) \mid x_1, \ldots, x_n)$ зависит не от самих выборочных значений, а только от функции $g(x_1, \ldots, x_n)$. Действительно, по формуле Байеса (см. (2.41) в первой книге)

$$W_1(\vartheta \mid x_1, \ldots, x_n) = \frac{w_1(\vartheta)\, W_n(x_1, \ldots, x_n \mid \vartheta)}{\int\limits_{-\infty}^{\infty} w_1(\vartheta)\, W_n(x_1, \ldots, x_n \mid \vartheta)\, d\vartheta},$$

и, подставляя выражение для функции правдоподобия из (2.28), получим *)

$$W_1(\vartheta \mid x_1, \ldots, x_n) = \frac{w_1(\vartheta)\, f[g(x_1, \ldots, x_n), \vartheta]}{\int\limits_{-\infty}^{\infty} w_1(\vartheta)\, f[g(x_1, \ldots, x_n), \vartheta]\, d\vartheta},$$

т. е.

$$W_1(\vartheta \mid x_1, \ldots, x_n) = W_1[\vartheta \mid g(x_1, \ldots, x_n)]. \quad (2.29)$$

Пусть $g_1(x_1, \ldots, x_n)$ и $g_2(x_1, \ldots, x_n)$ — две оценки параметра ϑ, определенные по выборке заданного размера n.

---

*) Можно также показать, что количество информации $I(x_1, \ldots, x_n, \vartheta)$, содержащееся в выборке $x_1, \ldots, x_n$ относительно неизвестного параметра ϑ, равно количеству информации $I[g(x_1, \ldots, x_n), \vartheta]$, содержащемуся в достаточной оценке $g(x_1, \ldots, x_n)$ этого параметра (см. [8, стр. 42—44]).

Оценку $g_1(x_1, \ldots, x_n)$ называют подчиненной оценке $g_2(x_1, \ldots, x_n)$, если из

$$g_2(x_1', \ldots, x_n') = g_2(x_1'', \ldots, x_n'')$$

следует:

$$g_1(x_1', \ldots, x_n') = g_1(x_1'', \ldots, x_n'').$$

Всякая оценка, подчиняющая достаточную статистику, сама является достаточной оценкой.

Если ϑ_0 — какое-нибудь фиксированное значение параметра ϑ, то можно доказать (см., например, [13]), что *отношение правдоподобия*

$$l(x_1, \ldots, x_n \mid \vartheta) = \frac{W_n(x_1, \ldots, x_n \mid \vartheta)}{W_n(x_1, \ldots, x_n \mid \vartheta_0)}$$

представляет достаточную статистику, из которой, используя приведенное выше замечание, можно найти достаточную оценку параметра ϑ.

Как показал Е. Б. Дынкин, для одномерных плотностей вероятностей, которые представимы в виде

$$w_1(x; \vartheta) = \exp\left\{\sum_{i=1}^{r} \varphi_i(x)\, c_i(\vartheta) + c_0(\vartheta) + \varphi_0(x)\right\},$$

существуют достаточные статистики параметра ϑ. Величину r называют рангом распределения. Функции $(1, \varphi_1, \ldots, \varphi_r)$ и $(1, c_1, \ldots, c_r)$ линейно независимы.

2.2.5. Эффективность. Условная оценка $\hat{\vartheta}_n$ называется эффективной оценкой параметра ϑ, если (при заданном смещении) среднее значение (по совокупности различных выборок заданного размера n) квадрата отклонения оценки от оцениваемого параметра не больше, чем среднеквадратическое отклонение для любой другой оценки, т. е.

$$m_1\{(\hat{\vartheta}_{n\,\text{эф}} - \vartheta)^2\} \leqslant m_1\{(\hat{\vartheta}_n - \vartheta)^2\}. \tag{2.30}$$

Иначе говоря, эффективная оценка (если она существует) минимизирует при заданном n величину $m_1\{(\hat{\vartheta}_n - \vartheta)^2\}$ по всем возможным условным оценкам $\hat{\vartheta}_n$ (с заданным смещением).

Если $\hat{\vartheta}_n$ — несмещенная оценка, то в (2.30) можно вместо ϑ подставить $m_1\{\hat{\vartheta}_n\}$ и переписать это неравенство в виде

$$M_2\{\hat{\vartheta}_{n\,\text{эф}}\} \leqslant M_2\{\hat{\vartheta}_n\}. \tag{2.31}$$

8—197

Следовательно, эффективная несмещенная оценка определяется из условия *минимума дисперсии* оценки.

Относительной эффективностью условной оценки $\hat{\vartheta}_n$ параметра ϑ называют отношение среднеквадратического отклонения эффективной оценки от оцениваемого параметра к среднеквадратическому отклонению рассматриваемой оценки (для смещенных оценок)

$$e_n(\vartheta) = \frac{m_1\{(\hat{\vartheta}_{n\ \text{эф}} - \vartheta)^2\}}{m_1\{(\hat{\vartheta}_n - \vartheta)^2\}} \tag{2.32}$$

и отношение дисперсии эффективной оценки к дисперсии рассматриваемой оценки (для несмещенных оценок)

$$e_n(\vartheta) = \frac{M_2\{\vartheta_{n\ \text{эф}}\}}{M_2\{\hat{\vartheta}_n\}}\ . \tag{2.33}$$

Ясно, что $0 \leqslant e_n \leqslant 1$. Величина относительной эффективности равна единице, когда оценка эффективна.

Иногда говорят об относительной эффективности одной оценки $\hat{\vartheta}_n^{(1)}$ параметра ϑ по отношению к другой оценке $\hat{\vartheta}_n^{(2)}$ этого же параметра, понимая под этим отношение среднеквадратических отклонений оценок от оцениваемого параметра. При таком определении относительная эффективность может быть любой неотрицательной величиной.

Следует отличать эффективную оценку от асимптотически эффективной. Оценка называется *асимптотически* эффективной, если предельное значение ее относительной эффективности, когда размер выборки n неограниченно увеличивается, отличается от нуля. Если среднеквадратическое отклонение оценки стремится к нулю при $n \to \infty$ (что обеспечивает состоятельность оценки), то эффективность означает, что при достаточно большом n среднеквадратические отклонения от оцениваемого параметра рассматриваемой и эффективной оценок представляют малые величины одного порядка.

Обратим внимание на то, что для смещенной оценки ее эффективность определяется не дисперсией, а среднеквадратическим отклонением от оцениваемого параметра. Легко указать тривиальный пример смещенной оценки с нулевой дисперсией. Пусть $\hat{\vartheta}_n = c_0 = \text{const}$ независимо от результатов наблюдений. Тогда $M_2\{\hat{\vartheta}_n\} = 0$. Но если только

величина оцениваемого параметра не угадана или не была заранее известна, смещение $b_n(\vartheta) = c_0 - \vartheta$ будет велико. Нельзя, вообще говоря, сделать равными нулю и дисперсию, и смещение оценки. Поэтому в дальнейшем случай нулевой дисперсии (которому соответствует $b'_n(\vartheta) = -1$) нами исключается.

Существует неравенство (известное как неравенство Рао — Крамера), с помощью которого можно определить нижнюю границу среднеквадратических отклонений оценок и тем самым оценить эффективность той или иной оценки.

Предположим, что границы области действительной оси x, где функция распределения $w_1(x; \vartheta)$ отлична от нуля, не зависят от ϑ. Это условие выполняется, например если $w_1(x; \vartheta) \neq 0$ на всей действительной оси или для $x > 0$. Примером распределения, которое не удовлетворяет этому условию, является равномерное

$$w_1(x; \vartheta) = \frac{1}{\vartheta}, \quad 0 < x < \vartheta.$$

Предположим, кроме того, что функция $w_1(x; \vartheta)$ дифференцируема по параметру ϑ. Пусть $\hat{\vartheta}_n = g(x_1, \ldots, x_n)$ — некоторая условная оценка параметра ϑ, полученная по выборке размером n. Среднее значение этой оценки

$$m_1\{\hat{\vartheta}_n\} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) W_n(x_1, \ldots \\ \ldots, x_n \mid \vartheta)\, dx_1 \ldots dx_n. \tag{2.34}$$

Используя (2.20), перепишем (2.34) в виде

$$\vartheta + b_n(\vartheta) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g(x_1, \ldots, x_n) W_n(x_1, \ldots \\ \ldots, x_n \mid \vartheta)\, dx_1, \ldots dx_n. \tag{2.35}$$

Предположим, что интеграл в правой части (2.35) можно дифференцировать по параметру ϑ. Тогда дифференцируя обе части (2.35) по ϑ и используя предположение о неза-

висимости пределов интегрирования от $\hat{\vartheta}$, получим

$$1+b_n'(\vartheta)=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} g(x_1,\ldots,x_n)\frac{\partial}{\partial\vartheta}W_n(x_1,\ldots,x_n|\vartheta)\times$$
$$\times dx_1\ldots dx_n=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} g(x_1,\ldots,x_n)\times$$
$$\times W_n(x_1,\ldots,x_n|\vartheta)\frac{\partial}{\partial\vartheta}\ln W_n(x_1,\ldots,x_n|\vartheta)\,dx_1\ldots dx_n=$$
$$=m_1\left\{\hat{\vartheta}_n\frac{\partial}{\partial\vartheta}\ln W_n\right\}. \qquad (2.36)$$

Кроме того, из очевидного равенства

$$\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} W_n(x_1,\ldots,x_n|\vartheta)\,dx_1\ldots dx_n=1$$

дифференцированием по ϑ находим

$$\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\frac{\partial}{\partial\vartheta}W_n(x_1,\ldots,x_n|\vartheta)\,dx_1\ldots dx_n=$$
$$=\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} W_n(x_1,\ldots,x_n|\vartheta)\frac{\partial}{\partial\vartheta}\ln W_n(x_1,\ldots$$
$$\ldots,x_n|\vartheta)\,dx_1\ldots dx_n=m_1\left\{\frac{\partial}{\partial\vartheta}\ln W_n\right\}=0. \qquad (2.37)$$

Умножая (2.37) на $m_1\{\hat{\vartheta}_n\}$ и вычитая из (2.36), получим

$$m_1\left\{[\hat{\vartheta}_n-m_1\{\hat{\vartheta}_n\}]\frac{\partial}{\partial\vartheta}\ln W_n\right\}=1+b_n'(\vartheta). \qquad (2.38)$$

Левая часть (2.38) представляет ковариацию двух случайных величин, имеющих нулевые средние. Как известно (см. стр. 84 в первой книге), квадрат этой ковариации не может превосходить произведение дисперсий сомножителей, т. е.

$$[1+b_n'(\vartheta)]^2\leqslant M_2\{\hat{\vartheta}_n\}\,M_2\left\{\frac{\partial}{\partial\vartheta}\ln W_n\right\}. \qquad (2.39)$$

Неотрицательная величина $M_2\left\{\frac{\partial}{\partial\vartheta}\ln W_n\right\}=$ $=m_1\left\{\left[\frac{\partial}{\partial\vartheta}\ln W_n\right]^2\right\}$ называется *информацией*, содержащейся

в выборке (по Р. А. Фишеру). Обозначим эту величину, зависящую только от размера выборки и функции распределения $w_1(x; \vartheta)$, через $I_n(\vartheta)$ и предположим, что она отлична от нуля *). Тогда из (2.39) находим искомую нижнюю границу дисперсии оценок **)

$$M_2\{\hat{\vartheta}_n\} \geqslant \frac{[1 + b'_n(\vartheta)]^2}{I_n(\vartheta)}. \qquad (2.40)$$

Заметим, что правая часть неравенства (2.40) является также нижней границей среднеквадратических отклонений оценок от оцениваемого параметра. Так как минимум величины $m_1\{[\hat{\vartheta}_n - \vartheta]^2\}$ достигается при $m_1\{\hat{\vartheta}_n\} = \vartheta$ (см. задачу 2.6 в первой книге), то

$$m_1\{[\hat{\vartheta}_n - \vartheta]^2\} \geqslant M_2\{\hat{\vartheta}_n\} \geqslant \frac{[1 + b'_n(\vartheta)]^2}{I_n(\vartheta)}. \qquad (2.41)$$

Для оценок с постоянным (не зависящим от ϑ) смещением ($b'(\vartheta) \equiv 0$), в частности и для несмещенных оценок ($b(\vartheta) \equiv 0$),

$$M_2\{\hat{\vartheta}_n\} \geqslant \frac{1}{I_n(\vartheta)}. \qquad (2.42)$$

В этом случае нижней границей дисперсии оценок является величина, обратная информации, содержащейся в выборке. Знаки равенства в (2.40) — (2.42) достигаются для эффективных оценок.

При ограниченном размере выборки случаи, когда в (2.42) достигается равенство, довольно редки. Как правило, нижняя грань дисперсий несмещенных оценок оказывается больше величины, обратной информации по Фишеру. Поэтому эффективность представляет не меру качества оценки, а меру качества нижней границы в неравенстве Рао — Крамера. Если дисперсия оценки достигает точной нижней границы, не равной $I_n^{-1}(\vartheta)$, то отличие эффективности от единицы нельзя считать показателем неудовлетворительности оценки (подробнее см. [2, гл. VIII], а также [11]).

---

*) Равенство нулю информации имеет место, например если функция правдоподобия не зависит от ϑ.

**) Формулу (2.40) называют неравенством Рао — Крамера, а иногда неравенством Фреше. Все три указанных автора пришли к этому результату независимо.

Нетрудно, используя (2.10), выразить информацию через функцию распределения $w_1(x;\ \vartheta)$. Так как

$$\ln W_n(x_1, \dots x_n \mid \vartheta) = \sum_{i=1}^{n} \ln w_1(x_i;\ \vartheta)$$

и

$$\frac{\partial}{\partial\vartheta} \ln W_n(x_1, \dots, x_n \mid \vartheta) = \sum_{i=1}^{n} \frac{1}{w_1(x_i;\ \vartheta)} \frac{\partial}{\partial\vartheta} w_1(x_i;\ \vartheta),$$

то с учетом (2.37)

$$I_n(\vartheta) = M_2\left\{\frac{\partial}{\partial\vartheta} \ln W_n\right\} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left(\frac{\partial}{\partial\vartheta} \ln W_n\right)^2 \times$$

$$\times W_n\, dx_1 \dots dx_n = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \Bigl[\sum_{i=1}^{n} \frac{1}{w_1(x_i;\vartheta)} \times$$

$$\times \frac{\partial w_1(x_i;\ \vartheta)}{\partial\vartheta}\Bigr]^2 \prod_{i=1}^{n} w_1(x_i;\vartheta)\, dx_1 \dots dx_n =$$

$$= \sum_{i=1}^{n} \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial\vartheta} \ln w_1(x_i;\vartheta)\right]^2 w_1(x_i;\ \vartheta)\, dx_i +$$

$$+ \sum_{i,\ j(i\neq j)}^{n} \int_{-\infty}^{\infty} \frac{\partial w_1(x_i;\vartheta)}{\partial\vartheta} dx_i \int_{-\infty}^{\infty} \frac{\partial w_1(x_j;\vartheta)}{\partial\vartheta} dx_j.$$

Так как слагаемые в первой сумме одинаковы и

$$\int_{-\infty}^{\infty} \frac{\partial w_1(x_k;\ \vartheta)}{\partial\vartheta} dx_k = \frac{\partial}{\partial\vartheta} \int_{-\infty}^{\infty} w_1(x_k;\ \vartheta)\, dx_k = 0,$$

то

$$I_n(\vartheta) = n \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial\vartheta} \ln w_1(x;\ \vartheta)\right]^2 w_1(x;\vartheta)\, dx =$$

$$= n m_1 \left\{\left[\frac{\partial}{\partial\vartheta} \ln w_1(x;\vartheta)\right]^2\right\} \qquad (2.43)$$

или

$$I_n(\vartheta) = nI_1(\vartheta). \qquad (2.43')$$

Таким образом, информация пропорциональна размеру выборки с коэффициентом пропорциональности, зависящим от исходного распределения.

Подставляя (2.43) в (2.41), получим

$$m_1\{[\hat{\vartheta}_n - \vartheta]^2\} \geqslant M_2\{\hat{\vartheta}_n\} \geqslant$$
$$\geqslant \frac{1}{n} \frac{[1 + b'_n(\vartheta)]^2}{\int\limits_{-\infty}^{\infty} \left[\frac{\partial}{\partial\vartheta} \ln w_1(x;\ \vartheta)\right]^2 w_1(x;\ \vartheta)\, dx}. \qquad (2.44)$$

Величину $m_1\{[\hat{\vartheta}_n - \vartheta]^2\}]^{1/2}$ принимают иногда в качестве меры *точности* оценки. Правая часть неравенства (2.44) определяет *потенциальную точность*.

Заметим, что для дискретного распределения выражение (2.43) преобразуется к виду

$$I_n(\vartheta) = n \sum_i \left[\frac{\partial \ln p(x_i;\ \vartheta)}{\partial\vartheta}\right]^2 p(x_i;\ \vartheta). \qquad (2.45)$$

Предположим теперь, что оценка $\hat{\vartheta}_n = g(x_1, \ldots, x_n)$ достаточная. Тогда из (2.28) следует:

$$\frac{\partial}{\partial\vartheta} \ln W_n(x_1, \ldots, x_n \mid \vartheta) = \frac{\partial}{\partial\vartheta} \ln f(\hat{\vartheta}_n,\ \vartheta).$$

Пусть, кроме того,

$$f(\hat{\vartheta}_n, \vartheta) = \exp\{k(\vartheta)[\hat{\vartheta}_n + c(\vartheta)]\} \qquad (2.46)$$

и, следовательно,

$$\frac{\partial}{\partial\vartheta} \ln f(\hat{\vartheta}_n, \vartheta) = k'(\vartheta)\, \hat{\vartheta}_n + \frac{d}{d\vartheta}[k(\vartheta)\, c(\vartheta)].$$

Тогда из (2.37) и (2.46) следует, что

$$m_1\left\{\frac{\partial}{\partial\vartheta} \ln f(\hat{\vartheta}_n,\ \vartheta\right\} = k'(\vartheta)\, m_1\{\hat{\vartheta}_n\} + \frac{d}{d\vartheta}[k(\vartheta)\, c(\vartheta)] = 0$$

и

$$\frac{\partial}{\partial\vartheta} \ln W_n = \frac{\partial}{\partial\vartheta} \ln f(\hat{\vartheta}_n, \vartheta) = k'(\vartheta)[\hat{\vartheta}_n - m_1\{\hat{\vartheta}_n\}]. \qquad (2.46')$$

Подставляя (2.46′) в (2.38), получим

$$k'(\vartheta)\, m_1\{[\hat{\vartheta}_n - m_1\{\hat{\vartheta}_n\}]^2\} = k'(\vartheta)\, M_2\{\hat{\vartheta}_n\} = 1 + b'_n(\vartheta). \ (2.47)$$

С другой стороны [см. (2.46′)],

$$I_n(\vartheta) = m_1\left\{\left(\frac{\partial}{\partial\vartheta}\ln W_n\right)^2\right\} = [k'(\vartheta)]^2 M_2\{\hat{\vartheta}_n\}. \quad (2.47')$$

Из (2.47) и (2.47′) находим

$$k'(\vartheta) = \frac{I_n(\vartheta)}{1 + b_n'(\vartheta)}, \quad (2.47'')$$

и, следовательно,

$$M_2\{\hat{\vartheta}_n\} = \frac{[1 + b_n'(\vartheta)]^2}{I_n(\vartheta)}.$$

Итак, доказано, что среди всех оценок с заданным смещением $b_n(\vartheta)$ *достаточная оценка определенного вида* (удовлетворяющая условиям (2.46) и (2.47″)) имеет минимальную дисперсию по сравнению с любыми другими оценками (см. (2.41)).

Таким образом, любая эффективная оценка принадлежит классу достаточных статистик, но лишь некоторые из достаточных оценок эффективны.

До сих пор в этом разделе рассматривались только условные оценки. Понятие эффективности можно распространить и на безусловные оценки, усредняя (2.30) по ϑ.

Неравенство Рао — Крамера (2.40) в этом случае перепишется в виде

$$\int_{-\infty}^{\infty} w_1(\vartheta)\, m_1\{[\hat{\vartheta}_n - \vartheta]^2\}\, d\vartheta \geqslant \frac{\int_{-\infty}^{\infty} w_1(\vartheta)\,[1 + b_n'(\vartheta)]^2\, d\vartheta}{\int_{-\infty}^{\infty} w_1(\vartheta)\, I_n(\vartheta)\, d\vartheta}.$$

2.2.6. Интервальные оценки. Под интервальной оценкой параметра ϑ понимают интервал, границы которого $\hat{\vartheta}_n^{(н)}$ и $\hat{\vartheta}_n^{(в)}$ являются функциями выборочных значений $x_1, \ldots, x_n$ и который содержит с заданной вероятностью оцениваемый параметр. Аналитически это может быть записано в виде

$$P\{\hat{\vartheta}_n^{(н)} < \vartheta < \hat{\vartheta}_n^{(в)}\} = \gamma. \quad (2.48)$$

Вероятность γ называется *коэффициентом доверия*, а оценки $\hat{\vartheta}_n^{(н)}$ и $\hat{\vartheta}_n^{(в)}$ — соответственно нижним и верхним *доверительными пределами*. Интервал $(\hat{\vartheta}_n^{(н)}, \hat{\vartheta}_n^{(в)})$ называется *доверительным*.

Иногда доверительный интервал строят следующим образом:

$$\hat{\vartheta}_n^{(\text{н})} = \hat{\vartheta}_n - \varepsilon_1 \vartheta, \tag{2.49}$$

$$\hat{\vartheta}_n^{(\text{в})} = \hat{\vartheta}_n + \varepsilon_2 \vartheta, \tag{2.49'}$$

где $\hat{\vartheta}_n$ — точечная оценка параметра ϑ; ε_1, ε_2 — положительные числа. Для заданного γ величины ε_1 и ε_2 можно найти бесконечным числом способов. Если $W_{\hat{\vartheta}_n}(z)$ функция распределения точечной оценки, то из (2.48), (2.49) и (2.49′) следует:

$$P\{\vartheta(1-\varepsilon_2) < \hat{\vartheta}_n < \vartheta(1+\varepsilon_1)\} = \gamma, \tag{2.50}$$

откуда получаем два соотношения для определения ε_1 и ε_2:

$$\int\limits_{-\infty}^{\vartheta(1-\varepsilon_2)} W_{\hat{\vartheta}_n}(z)\,dz = \gamma_2, \tag{2.51}$$

$$\int\limits_{\vartheta(1+\varepsilon_1)}^{\infty} W_{\hat{\vartheta}_n}(z)\,dz = \gamma_1, \tag{2.51'}$$

где γ_1, γ_2 — любые положительные числа, меньшие единицы, причем $\gamma_1 + \gamma_2 = 1 - \gamma$.

При $\varepsilon_1 = \varepsilon_2 = \varepsilon$ формула (2.50) определяет связь между коэффициентом доверия γ, относительной длиной доверительного интервала 2ε и размером выборки n:

$$P\{\vartheta(1-\varepsilon) < \hat{\vartheta}_n < \vartheta(1+\varepsilon)\} = \int\limits_{\vartheta(1-\varepsilon)}^{\vartheta(1+\varepsilon)} W_{\hat{\vartheta}_n}(z)\,dz = \gamma. \tag{2.52}$$

Если задана длина доверительного интервала, то для состоятельных и несмещенных оценок коэффициент доверия будет возрастать по мере увеличения размера выборки, приближаясь к единице. При заданном размере выборки коэффициент доверия будет тем больше, чем шире доверительный интервал. Иными словами, при заданном размере выборки невозможно повысить коэффициент доверия, не увеличивая длину доверительного интервала, или нельзя уменьшить этот интервал, не уменьшая коэффициента доверия.

Возможны три вида задач, использующих интервальные оценки параметра. Для выборки заданного размера n строится точечная оценка $\hat{\vartheta}_n$, находится ее распределение

$W\hat{\vartheta}_n(z)$ и для фиксированного ε из (2.52) определяется коэффициент доверия γ. При тех же условиях можно по заданному γ найти относительную длину доверительного интервала 2ε. Наконец, можно задать и коэффициент доверия γ, и относительную длину доверительного интервала 2ε. Тогда из (2.52) последовательными приближениями находят размер выборки, для которой возможно достигнуть одновременно заданных γ и ε.

2.3. МЕТОДЫ ПОЛУЧЕНИЯ ОЦЕНОК ПО ОПРЕДЕЛЕННЫМ КРИТЕРИЯМ

2.3.1. Оценка максимального правдоподобия. Функция правдоподобия выборки размером n для заданных выборочных значений $x_1, \ldots, x_n$ является функцией неизвестного параметра ϑ [см. (2.10)]

$$W_n(x_1, \ldots, x_n \mid \vartheta) = \prod_{k=1}^{n} w_1(x_k; \vartheta).$$

Чтобы подчеркнуть это обстоятельство, введем новое обозначение функции правдоподобия:

$$L_x(\vartheta) = W_n(x_1, \ldots, x_n \mid \vartheta). \qquad (2.53)$$

Выберем в качестве условной оценки неизвестного параметра то значение переменной ϑ, для которого при заданных $x_1, \ldots, x_n$ функция правдоподобия $L_x(\vartheta)$ достигает максимума. Оценка, удовлетворяющая указанному критерию качества, называется *максимально правдоподобной*.

Условие экстремума функции правдоподобия записывается в виде

$$\frac{\partial L_x(\vartheta)}{\partial \vartheta} = 0. \qquad (2.54)$$

Практически, однако, уравнение (2.54) не используется. Простой математический прием упрощает представление этого уравнения посредством исходного распределения $w_1(x; \vartheta)$. Так как логарифм — монотонная функция, то экстремумы функций $L_x(\vartheta)$ и $\ln L_x(\vartheta)$ достигаются при одинаковых значениях аргумента ϑ. Поэтому уравнение правдоподобия записывают обычно в виде

$$\frac{\partial \ln L_x(\vartheta)}{\partial \vartheta} = 0 \qquad (2.55)$$

или с учетом (2.10)

$$\sum_{k=1}^{n} \frac{\partial \ln w_1(x_k;\,\vartheta)}{\partial \vartheta} = 0. \qquad (2.56)$$

Для дискретного распределения с учетом (2.10′) получаем

$$\sum_{k=1}^{n} \frac{\partial \ln p\,(x_k;\,\vartheta)}{\partial \vartheta} = 0. \qquad (2.56')$$

Уравнение правдоподобия представляет в общем случае нелинейное алгебраическое или трансцендентное уравнение, которое может иметь много решений, соответствующих относительным максимумам и минимумам функции правдоподобия. Каждое решение $\hat{\vartheta}_n = g\,(x_1, \ldots, x_n)$, соответствующее максимуму функции правдоподобия, представляет оценку максимального правдоподобия неизвестного параметра. Задача в этом случае состоит в нахождении того решения, которое соответствует абсолютному максимуму (максимуму максиморуму) функции правдоподобия.

Если для заданного размера выборки n существует несмещенная эффективная условная оценка $\hat{\vartheta}_{n\,\text{эф}}$ параметра ϑ, то уравнение правдоподобия (2.55) в соответствии с (2.46′) может быть записано в виде

$$\frac{\partial \ln L_x(\vartheta)}{\partial \vartheta} = k'(\vartheta)\,[\hat{\vartheta}_{n\,\text{эф}} - \vartheta] = 0,$$

где согласно (2.47″)

$$k'(\vartheta) = I_n(\vartheta) > 0.$$

Так как $\frac{\partial \ln L_x(\vartheta)}{\partial \vartheta} > 0$ при $\vartheta < \hat{\vartheta}_{n\,\text{эф}}$ и $\frac{\partial \ln L_x(\vartheta)}{\partial \vartheta} < 0$ при $\vartheta > \hat{\vartheta}_{n\,\text{эф}}$, то отсюда следует, что оценка максимального правдоподобия равна

$$\hat{\vartheta}_n = \hat{\vartheta}_{n\,\text{эф}}.$$

Таким образом приходим к следующему важному заключению.

Если существует несмещенная эффективная оценка $\hat{\vartheta}_{n\,\text{эф}}$, то уравнение правдоподобия имеет *единственное* решение, равное $\hat{\vartheta}_{n\,\text{эф}}$.

Если существует достаточная оценка $\hat{\vartheta}_n = g(x_1, \ldots, x_n)$, то из (2.28) и (2.55) следует, что уравнение правдоподобия может быть записано в виде

$$\frac{\partial \ln L_x(\vartheta)}{\partial \vartheta} = \frac{\partial \ln f[g(x_1, \ldots, x_n), \vartheta]}{\partial \vartheta} = 0.$$

Следовательно, если существует достаточная оценка, то каждое решение уравнения правдоподобия является функцией этой достаточной оценки.

Доказано (см. [5, § 33]), что при некоторых не очень ограничительных предположениях относительно функции распределения $w_1, (x; \vartheta)$ оценка максимального правдоподобия *состоятельная*. Эта оценка *асимптотически несмещенная и асимптотически эффективная*. Кроме того, при $n \to \infty$ распределение оценки максимального правдоподобия — асимптотически нормальное со средним значением ϑ и дисперсией, равной обратной величине информации [см. (2.48)], т. е.

$$W_{\hat{\vartheta}_n}(z) \sim \sqrt{\frac{I_n(\vartheta)}{2\pi}} \exp\left[-\frac{I_n(\vartheta)}{2}(z-\vartheta)^2\right], \qquad (2.57)$$

причем $I_n(\nu) \to \infty$ при $n \to \infty$.

Подчеркнем, что оценка максимального правдоподобия лишь асимптотически несмещенная и эффективная. При конечном объеме выборки могут, вообще говоря, иметь место и смещение, и существенное отличие дисперсии оценки от нижней границы дисперсии.

2.3.2. Приближенное решение уравнения правдоподобия. Точного решения уравнения правдоподобия (2.55) не удается получить, как правило, простыми приемами. В общем случае для получения решения может быть рекомендован следующий метод последовательных приближений *).

Разложим логарифм функции правдоподобия в ряд Тейлора около точки ϑ_0 и ограничимся в этом разложении первыми тремя членами:

$$\ln L_x(\vartheta) \approx \ln L_x(\vartheta_0) + (\vartheta - \vartheta_0)\left[\frac{\partial}{\partial \vartheta} \ln L_x(\vartheta)\right]_{\vartheta=\vartheta_0} + \frac{1}{2}(\vartheta-\vartheta_0)^2\left[\frac{\partial^2}{\partial \vartheta^2} \ln L_x(\vartheta)\right]_{\vartheta=\vartheta_0}. \qquad (2.58)$$

---

*) Вариант так называемого итерационного метода Ньютона. Подробнее см., например: Бут Э. Д. Численные методы. Физматгиз, 1959, стр. 187.

Из (2.58) непосредственно находим

$$\frac{\partial \ln L_x(\vartheta)}{\partial \vartheta} \approx \left[\frac{\partial}{\partial \vartheta} \ln L_x(\vartheta)\right]_{\vartheta=\vartheta_0} + \\ + (\vartheta - \vartheta_0)\left[\frac{\partial^2}{\partial \vartheta^2} \ln L_x(\vartheta)\right]_{\vartheta=\vartheta_0}. \quad (2.59)$$

Примем в качестве первого приближения решения уравнения правдоподобия *) какую-нибудь грубую оценку $\hat{\vartheta}_1$ неизвестного параметра ϑ. Это может быть выборочное среднее, или выборочная медиана, или какая-либо иная оценка. Используем теперь (2.59) для определения второго приближения $\hat{\vartheta}_2$ как корня уравнения

$$\left[\frac{\partial}{\partial \vartheta} \ln L_x(\vartheta)\right]_{\vartheta=\hat{\vartheta}_1} + (\vartheta - \hat{\vartheta}_1)\left[\frac{\partial^2}{\partial \vartheta^2} \ln L_x(\vartheta)\right]_{\vartheta=\hat{\vartheta}_1} = 0. \quad (2.60)$$

Из (2.60) следует:

$$\hat{\vartheta}_2 = \hat{\vartheta}_1 - \frac{\left[\frac{\partial}{\partial \vartheta} \ln L_x(\vartheta)\right]_{\vartheta=\hat{\vartheta}_1}}{\left[\frac{\partial^2}{\partial \vartheta^2} \ln L_x(\vartheta)\right]_{\vartheta=\hat{\vartheta}_1}}. \quad (2.61)$$

Заменим $\frac{\partial^2}{\partial \vartheta^2} \ln L_x(\vartheta)$ его средним значением (по пространству выборок):

$$\frac{\partial^2}{\partial \vartheta^2} \ln L_x(\vartheta) \approx m_1\left\{\frac{\partial^2}{\partial \vartheta^2} \ln L_x(\vartheta)\right\} = \\ = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\partial^2}{\partial \vartheta^2}[\ln L_x(\vartheta)]\, L_x(\vartheta)\, dx_1 \ldots dx_n = \\ = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{\frac{\partial^2 L_x(\vartheta)}{\partial \vartheta^2} - \left[\frac{1}{L_x(\vartheta)} \frac{\partial L_x(\vartheta)}{\partial \vartheta}\right]^2 L_x(\vartheta)\right\} dx_1 \ldots dx_n. \quad (2.62)$$

В силу (2.37)

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\partial^2 L_x(\vartheta)}{\partial \vartheta^2}\, dx_1 \ldots dx_n = 0. \quad (2.63)$$

---

*) Здесь молчаливо предполагается, что функция правдоподобия унимодальна или что первое приближение находится вблизи главного максимума этой функции.

Тогда из (2.62) получаем

$$\frac{\partial^2}{\partial\vartheta^2}\ln L_x(\vartheta)\approx m_1\left\{\frac{\partial^2}{\partial\vartheta^2}\ln L_x(\vartheta)\right\}=$$
$$=-m_1\left\{\left[\frac{\partial\ln L_x(\vartheta)}{\partial\vartheta}\right]^2\right\}=-I_n(\vartheta). \qquad (2.64)$$

Подставляя (2.64) в (2.61), находим выражение для второго приближения:

$$\hat{\vartheta}_2=\hat{\vartheta}_1+\frac{1}{I_n(\hat{\vartheta}_1)}\left[\frac{\partial}{\partial\vartheta}\ln L_x(\vartheta)\right]_{\vartheta=\hat{\vartheta}_1}. \qquad (2.65)$$

Следующее приближение можно получить, подставляя в правую часть (2.59) вместо ϑ_0 величину $\hat{\vartheta}_2$ из (2.65), приравнивая ее нулю и находя корень уравнения с заменой в нем отрицательной второй производной ее средним значением — информацией $I_n(\hat{\vartheta}_2)$. Если найдено m-е приближение, то следующее приближение находится по формуле

$$\hat{\vartheta}_{m+1}=\hat{\vartheta}_m+\frac{1}{I_n(\hat{\vartheta}_m)}\left[\frac{\partial}{\partial\vartheta}\ln L_x(\vartheta)\right]_{\vartheta=\hat{\vartheta}_m}, \qquad (2.66)$$

являющейся очевидным обобщением (2.65).

2.3.3. Максимум апостериорной плотности вероятности оцениваемого параметра. Предположим теперь, что неизвестный параметр ϑ — случайная величина, функция распределения которой $w_1(\vartheta)$. По формуле Байеса (см. (2.41) в первой книге) можно определить условную плотность вероятности параметра для заданной выборки $x_1, \ldots, x_n$:

$$W_1(\vartheta\mid x_1, \ldots, x_n)=\frac{w_1(\vartheta)L_x(\vartheta)}{\int\limits_{-\infty}^{\infty}w_1(\vartheta)L_x(\vartheta)\,d\vartheta}. \qquad (2.67)$$

Выберем в качестве оценки параметра то значение переменной ϑ, для которого при заданных $x_1, \ldots, x_n$ апостериорная плотность вероятности оцениваемого параметра максимальна. Учитывая замечание, которое было сделано при выводе уравнения правдоподобия (2.55), можно уравнение для оценки, соответствующей максимальной апостериорной плотности вероятности, представить в виде

$$\frac{\partial}{\partial\vartheta}\ln[w_1(\vartheta)L_x(\vartheta)]=0 \qquad (2.68)$$

или

$$\frac{\partial}{\partial\vartheta}\ln w_1(\vartheta)+\sum_{k=1}^{n}\frac{\partial}{\partial\vartheta}\ln w_1(x_k\,|\,\vartheta)=0. \qquad (2.69)$$

Рассматриваемые оценки обладают свойствами, схожими со свойствами условных оценок максимального правдоподобия [9, § 21.1.2]. Если $\hat{\Theta}_{n\ \text{эф}}$ — несмещенная эффективная безусловная оценка параметра ϑ, то уравнение (2.68) имеет единственное решение, равное $\hat{\Theta}_{n\ \text{эф}}$. Оценка, соответствующая максимуму апостериорной плотности вероятности оцениваемого параметра, при слабых ограничениях состоятельна и асимптотически эффективна. Функция распределения этой оценки при $n\to\infty$ может быть представлена асимптотическим соотношением

$$W_{\hat{\theta}_n}(z)\sim\sqrt{\frac{m_1\{I_n(\vartheta)\}}{2\pi}}\exp\left\{-\frac{m_1\{I_n(\vartheta)\}}{2}[z-m_1\{\vartheta\}]^2\right\}. \qquad (2.70)$$

Если априорное распределение параметра ϑ точно или приблизительно равномерное на некотором конечном интервале, то из (2.67) следует:

$$W_1(\vartheta\,|\,x_1,\ \ldots,\ x_n)=kL_x(\vartheta),$$

где k — константа, не зависящая от ϑ.

В этом случае оценки максимального правдоподобия будут одновременно оценками, соответствующими максимумам апостериорной плотности вероятности оцениваемого параметра ϑ.

2.3.4. Байесовские оценки. Точечные оценки, полученные по критериям максимального правдоподобия или максимальной апостериорной плотности вероятности соответствуют требованиям специфического вида и могут недостаточно полно отражать потери, которые возникают в связи с плохими оценками. Действительно, с каждой оценкой $\hat{\vartheta}_n = g(x_1,\ \ldots,\ x_n)$, образованной на основании наблюдаемой выборки $x_1,\ \ldots,\ x_n$, связана *ошибка*, измеряемая разностью $\vartheta-\hat{\vartheta}_n$. Эта разность будет изменяться от выборки к выборке: в одном случае она может быть невелика, а в другом — большой. Для того чтобы учесть это обстоятельство, каждой ошибке приписывают определенный вес и вводят неотрица-

тельную *функцию потерь* $\Pi(\hat{\vartheta}_n, \vartheta)$. В общем случае эта функция может зависеть не только от разности $\vartheta - \hat{\vartheta}_n$, а от каждой из величин, указанных в ее аргументе.

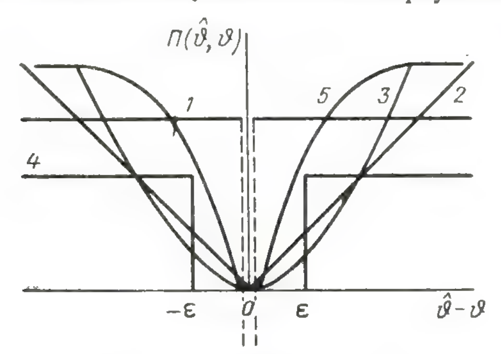


Рис. 14. Функции потерь:
1 — простая, 2 — модуль ошибки, 3 — квадратичная, 4 — прямоугольная, 5 — экспоненциальная.

Выбор вида функции потерь в известной степени субъективен и зависит от конкретной ситуации. Наиболее распространенными в приложениях функциями являются следующие (рис. 14):

простая функция потерь *)

$$\Pi(\hat{\vartheta}_n, \vartheta) = c - \delta(\hat{\vartheta}_n - \vartheta), \qquad c > 0; \tag{2.71}$$

функция потерь, равная модулю ошибки

$$\Pi(\hat{\vartheta}_n, \vartheta) = |\hat{\vartheta}_n - \vartheta|, \tag{2.72}$$

квадратичная функция потерь

$$\Pi(\hat{\vartheta}_n, \vartheta) = (\hat{\vartheta}_n - \vartheta)^2; \tag{2.73}$$

прямоугольная функция потерь

$$\Pi(\hat{\vartheta}_n, \vartheta) = \begin{cases} 0, & |\vartheta - \hat{\vartheta}_n| < \varepsilon, \\ 1, & |\vartheta - \hat{\vartheta}_n| \geq \varepsilon; \end{cases} \tag{2.74}$$

экспоненциальная функция потерь

$$\Pi(\hat{\vartheta}_n, \vartheta) = 1 - \exp[-k(\hat{\vartheta}_n - \vartheta)^2]. \tag{2.75}$$

---

*) Функция (2.71) неотрицательна, за исключением точки $\hat{\vartheta}_n = \vartheta$.

Все приведенные функции потерь являются *симметричными* функциями ошибки $\hat{\vartheta}_n - \vartheta$.

Мерилом качества условной оценки может служить среднее значение потерь по всевозможным выборкам, учитывающее частоту появления тех или иных выборок, т. е. функцию правдоподобия $L_x(\vartheta)$. Указанное среднее значение, равное

$$r(\vartheta) = m_1\{\Pi(\hat{\vartheta}_n, \vartheta)\} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \Pi(\hat{\vartheta}_n, \vartheta) L_x(\vartheta)\, dx_1 \ldots dx_n, \tag{2.76}$$

называется условной функцией риска. Потребуем теперь, чтобы условная оценка $\hat{\vartheta}_n$ минимизировала функционал (2.76). Оценка, оптимальная по критерию минимума условной функции риска, называется *условной байесовской оценкой.*

Если оцениваемый параметр ϑ — случайный и $w_1(\vartheta)$ — его функция распределения, то, усредняя условную функцию риска по всевозможным значениям ϑ, получаем так называемый *средний риск*

$$R = \int_{-\infty}^{\infty} w_1(\vartheta)\, r(\vartheta)\, d\vartheta. \tag{2.77}$$

Подставляя в (2.77) выражение для $r(\vartheta)$ из (2.76) и учитывая (2.67), представим средний риск в виде

$$R = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} J(\hat{\vartheta}_n \mid x_1, \ldots, x_n) W_n(x_1, \ldots, x_n)\, dx_1 \ldots, dx_n = m_1\{J(\hat{\vartheta}_n \mid x_1, \ldots, x_n)\}, \tag{2.78}$$

$$J(\hat{\vartheta}_n \mid x_1, \ldots, x_n) = \int_{-\infty}^{\infty} \Pi(\hat{\vartheta}_n, \vartheta) W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta; \tag{2.79}$$

$$W_n(x_1, \ldots, x_n) = \int_{-\infty}^{\infty} w_1(\vartheta) L_x(\vartheta)\, d\vartheta. \tag{2.79'}$$

Оценки, получаемые по критерию минимального среднего риска, называются *безусловными байесовскими оценками* (или просто байесовскими оценками).

Байесовская оценка при заданных функции потерь и априорном распределении параметра ϑ находится из условия минимума функционала (2.78), зависящего от вида функции $\hat{\vartheta}_n = g(x_1, \ldots, x_n)$. Так как функционал J зависит от g и не зависит от частных производных $\frac{\partial g}{\partial x_i}$, то известное в вариационном исчислении уравнение Эйлера для определения экстремальной функции запишется в виде

$$\frac{\partial J}{\partial g} = 0 \qquad (2.80)$$

при условии, что производная в левой части (2.80) существует.

2.3.5. Простая функция потерь. Запишем выражение условной функции риска при простой функции потерь. Подставляя (2.71) в (2.76) и используя фильтрующее свойство дельта-функции в многомерном интеграле (см. (3.117) в первой книге), получим

$$r(\vartheta) = c - L_{\hat{\vartheta}_n}(\vartheta), \qquad (2.81)$$

где $L_{\hat{\vartheta}_n}(\vartheta)$ — плотность вероятности оценки $\hat{\vartheta}_n$ при заданном ϑ. Для того чтобы $\hat{\vartheta}_n$ была условной байесовской оценкой $\hat{\vartheta}_n^*$, она должна минимизировать функцию риска $r(\vartheta)$, что эквивалентно требованию:

$$L_{\hat{\vartheta}_n^*}(\vartheta) \geqslant L_{\hat{\vartheta}_n}(\vartheta). \qquad (2.82)$$

Последнее означает, что $\hat{\vartheta}_n^*$ является условной оценкой максимального правдоподобия.

Таким образом, условные оценки максимального правдоподобия являются частным видом *условных байесовских оценок при простой функции потерь*.

Подставляя (2.81) в (2.77), находим также выражение среднего риска при простой функции потерь:

$$R = c - \int_{-\infty}^{\infty} w_1(\vartheta) L_{\hat{\vartheta}_n}(\vartheta)\, d\vartheta, \qquad (2.83)$$

из которого следует, что оценка $\hat{\vartheta}_n^*$ будет безусловной байесовской оценкой, если

$$w_1(\vartheta) L_{\hat{\vartheta}_n^*}(\vartheta) \geqslant w_1(\vartheta) L_{\hat{\vartheta}_n}(\vartheta). \qquad (2.84)$$

Таким образом, оценка, соответствующая максимуму апостериорной плотности вероятности [см. (2.67)], является частным видом *безусловных байесовских оценок при простой функции потерь.*

2.3.6. Квадратичная функция потерь. Условная функция риска в этом случае равна

$$r(\vartheta) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} (\hat{\vartheta}_n - \vartheta)^2 L_x(\vartheta)\, dx_1 \ldots dx_n, \quad (2.85)$$

т. е. совпадает со средним квадратическим отклонением оценки от оцениваемого параметра (или с дисперсией оценки, если она несмещенная). Поэтому условие минимума функции риска $r(\vartheta)$ сводится к минимизации $m_1\{(\hat{\vartheta}_n - \vartheta)^2\}$, т. е. к отысканию эффективной оценки (см. § 2.2.5). Таким образом, *эффективная оценка (условная) является условной байесовской оценкой при квадратичной функции потерь.* Подставляя (2.73) в (2.79), находим, что функционал $J(\hat{\vartheta}_n \mid x_1, \ldots, x_n)$ в выражении среднего риска (2.78) при квадратичной функции потерь имеет следующий вид:

$$J(\hat{\vartheta}_n \mid x_1, \ldots, x_n) = \int_{-\infty}^{\infty} (\hat{\vartheta}_n - \vartheta)^2 W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta. \quad (2.86)$$

Подставляя (2.86) в (2.80) и разрешая уравнение относительно неизвестной функции $\hat{\vartheta}_n = g(x_1, \ldots, x_n)$, получим искомое выражение байесовской оценки *):

$$\hat{\vartheta}_n^* = g^*(x_1, \ldots, x_n) = \int_{-\infty}^{\infty} \vartheta W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta \quad (2.87)$$

или

$$\hat{\vartheta}_n^* = m_1\{\vartheta \mid x_1, \ldots, x_n\} = \frac{\int_{-\infty}^{\infty} \vartheta w_1(\vartheta)\, L_x(\vartheta)\, d\vartheta}{\int_{-\infty}^{\infty} w_1(\vartheta)\, L_x(\vartheta)\, d\vartheta}. \quad (2.87')$$

Из (2.87) следует, что байесовская оценка при квадратичной функции потерь совпадает с *условным средним значением*

*) Нетрудно убедиться, что (2.87) соответствует *минимуму* среднего риска R.

оцениваемого параметра ϑ при заданных выборочных значениях $x_1, \ldots, x_n$.

Подставляя (2.87) в (2.86), находим (при несмещенной оценке)

$$J^* = J(\hat{\vartheta}_n^* \mid x_1, \ldots, x_n) = M_2\{\vartheta \mid x_1, \ldots, x_n\},$$

и, следовательно, минимальный (байесовский) риск при квадратичной функции потерь равен

$$R^* = m_1\{M_2\{\vartheta \mid x_1, \ldots, x_n\}\} = M_2\{\vartheta\}, \tag{2.88}$$

т. е. дисперсии оцениваемого случайного параметра ϑ.

В отличие от простой функции потерь, для которой байесовская оценка определяется локальными свойствами апостериорной плотности вероятности оцениваемого параметра ϑ в окрестности ее максимума, байесовская оценка при квадратичной функции потерь зависит от изменения указанной апостериорной плотности во всем диапазоне изменения параметра ϑ. Заметим, однако, что для унимодальной и симметричной относительно моды апостериорной функции распределения $W_1(\vartheta \mid x_1, \ldots, x_n)$ байесовская оценка при квадратичной функции потерь совпадает, как это следует из (2.87), с оценкой по критерию максимума апостериорной плотности, т. е. с байесовской оценкой при простой функции потерь.

2.3.7. Функция потерь, равная модулю ошибки. Для функции потерь (2.72) находим, что функция $J(\hat{\vartheta}_n \mid x_1, \ldots, x_n)$ в выражении среднего риска (2.78) имеет вид

$$J(\hat{\vartheta}_n \mid x_1, \ldots, x_n) = \int_{-\infty}^{\infty} |\hat{\vartheta}_n - \vartheta| W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta =$$
$$= \int_{-\infty}^{\hat{\vartheta}_n} (\hat{\vartheta}_n - \vartheta) W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta -$$
$$- \int_{\hat{\vartheta}_n}^{\infty} (\hat{\vartheta}_n - \vartheta) W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta. \tag{2.89}$$

Подставляя (2.89) в (2.80), получим

$$\frac{\partial J}{\partial \hat{\vartheta}_n} = \int_{-\infty}^{\hat{\vartheta}_n} W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta - \int_{\hat{\vartheta}_n}^{\infty} W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta = 0$$

или

$$\int\limits_{-\infty}^{\hat{\vartheta}_n} W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta = \int\limits_{\hat{\vartheta}_n}^{\infty} W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta. \qquad (2.89')$$

Из (2.89′) следует, что байесовская оценка при функции потерь, равной модулю ошибки, совпадает с *условной медианой* оцениваемого параметра ϑ при заданных выборочных значениях $x_1, \ldots, x_n$.

Если апостериорная плотность вероятности оцениваемого параметра имеет один максимум (унимодальна) и симметрична относительно моды, то медиана и среднее значение этого распределения совпадают и равны его моде. В этом случае байесовские оценки при функции потерь, равной модулю ошибки, и при квадратичной функции потерь одинаковы и совпадают с оценкой максимальной апостериорной вероятности.

2.3.8. Прямоугольная функция потерь. Для функции потерь (2.74) из (2.79) находим, что функция $J(\hat{\vartheta}_n \mid x_1, \ldots$ $\ldots, x_n)$ в выражении среднего риска (2.78) имеет вид

$$J(\hat{\vartheta}_n \mid x_1, \ldots, x_n) = \int\limits_{-\infty}^{\hat{\vartheta}_n - \varepsilon} W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta + \int\limits_{\hat{\vartheta}_n + \varepsilon}^{\infty} W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta = 1 - \int\limits_{\hat{\vartheta}_n - \varepsilon}^{\hat{\vartheta}_n + \varepsilon} W_1(\vartheta \mid x_1, \ldots, x_n)\, d\vartheta. \qquad (2.90)$$

Подставляя (2.90) в (2.80), получим следующее трансцендентное уравнение для определения байесовской оценки при прямоугольной функции потерь:

$$W_1(\hat{\vartheta}_n + \varepsilon \mid x_1, \ldots, x_n) = W_1(\hat{\vartheta}_n - \varepsilon \mid x_1, \ldots, x_n). \qquad (2.90')$$

Если апостериорная плотность вероятности оцениваемого параметра унимодальна и симметрична относительно моды, то единственным решением (2.90′) является такая оценка $\hat{\vartheta}_n$, которая совпадает с модой указанной апостериорной плотности вероятности. В этом случае, следовательно,

байесовская оценка при прямоугольной функции потерь совпадает с оценкой, соответствующей максимальной апостериорной вероятности (т. е. с байесовской оценкой при простой и квадратичной функциях потерь), условная байесовская оценка при этом является оценкой максимального правдоподобия.

2.3.9. Симметричная функция потерь. Рассмотрим произвольную дифференцируемую функцию потерь, четную относительно ошибки, т. е. удовлетворяющую условию

$$\Pi(\hat{\vartheta}_n - \vartheta) = \Pi(\vartheta - \hat{\vartheta}_n). \tag{2.91}$$

Предположим, что условная плотность вероятности параметра ϑ при заданных выборочных значениях $x_1, \ldots, x_n$ унимодальна и симметрична относительно моды. Из этого предположения следует, что условное среднее является модой условной плотности, т. е. $W_1(\vartheta - m_1\{\vartheta \mid x_1, \ldots, x_n\} \mid x_1, \ldots, x_n)$ — четная функция аргумента $\vartheta - m_1$. Запишем теперь выражение для производной функционала (2.79):

$$\frac{\partial J}{\partial g} = \int_{-\infty}^{\infty} \frac{\partial}{\partial g}[\Pi(g-\vartheta)] \times$$
$$\times W_1(\vartheta - m_1\{\vartheta \mid x_1, \ldots, x_n\} \mid x_1, \ldots, x_n)\, d\vartheta. \tag{2.92}$$

Так как $\Pi(g - \vartheta)$ — четная функция, то ее производная — нечетная функция аргумента $g - \vartheta$. Так как второй сомножитель подынтегральной функции (2.92) — четная функция $\vartheta - m_1\{\vartheta \mid x_1, \ldots, x_n\}$, то величина $\frac{\partial J}{\partial g}$ тождественно обращается в нуль, если оценка $\hat{\vartheta}_n = g(x_1, \ldots, x_n)$ равна

$$\hat{\vartheta}_n = m_1\{\vartheta \mid x_1, \ldots, x_n\}, \tag{2.93}$$

потому что при выполнении равенства (2.93) подынтегральная функция становится нечетной функцией относительно новой переменной интегрирования $\tau = \vartheta - m_1\{\vartheta \mid x_1, \ldots, x_n\}$.

Таким образом, оценка (2.93) является решением уравнения (2.80) и, следовательно, *байесовской оценкой*.

Сравнивая (2.93) с (2.87′), приходим к выводу, что байесовская оценка при квадратичной функции потерь (с произвольной апостериорной плотностью оцениваемого параметра) является также байесовской оценкой при сим-

метричной функции потерь, если апостериорная плотность оцениваемого параметра унимодальна и симметрична относительно моды.

2.3.10. Минимаксные оценки. Когда априорное распределение $w_1(\vartheta)$ неизвестно, процедура построения безусловной оценки случайного параметра ϑ может основываться на *минимаксном* критерии качества. Рассмотрим семейство

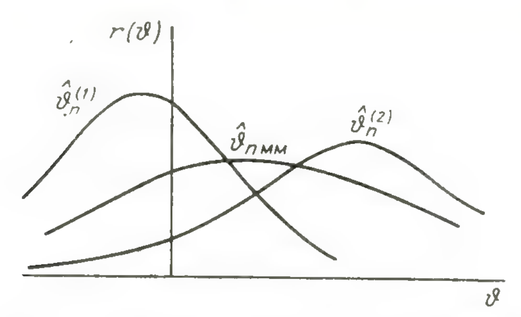


Рис. 15. Семейство кривых условных функций риска для различных оценок.

кривых условных функций риска, построенных для различных оценок $\hat{\vartheta}_n$ (рис. 15). Минимаксной называется оценка $\hat{\vartheta}_{n\,\text{мм}}$, для которой верхняя граница значений условной функции риска $r(\vartheta)$ не превосходит верхних границ значений функции (относительно переменной ϑ) при любых других оценках. Так же как и минимаксное правило выбора решения в альтернативных ситуациях, минимаксная оценка дает уверенность в том, что потери в среднем (по совокупности выборок заданного размера) не будут больше некоторой величины r_{min}. В некоторых случаях манимаксная оценка может оказаться слишком осторожной (см. рис. 1 на стр. 22).

Как доказал А. Вальд (см. стр. 21—22), минимаксная оценка является байесовской оценкой при наименее благоприятном априорном распределении оцениваемого параметра, для которого средний риск (при заданной функции потерь) имеет наибольшую величину. Он же показал, что байесовская оценка, для которой условная функция риска превращается в константу (т. е. не зависит от ϑ), представляет минимаксную оценку. Однако обращение этой функции в постоянную величину при небайесовской оценке не озна-

чает, что оценка минимаксная. С другой стороны, если условие $r(\vartheta) = \text{const}$ неосуществимо для всех ϑ при байесовских оценках, то это еще не значит, что нельзя найти минимаксную оценку.

Из (2.77) следует, что при $r(\vartheta) = r_{min}$

$$R_{max} = r_{min}, \tag{2.94}$$

т. е. максимальное значение среднего риска совпадает с минимальным значением функции риска, если минимаксная оценка является байесовской.

2.4. СОВМЕСТНЫЕ ОЦЕНКИ СОВОКУПНОСТИ ПАРАМЕТРОВ

2.4.1. Основные свойства. Изложенную выше теорию оценок можно распространить на те случаи, когда исходная функция распределения зависит от нескольких неизвестных параметров $\vartheta_1, \ldots, \vartheta_m$. По выборке заданного размера n из распределения $w_1(x; \vartheta_1, \ldots, \vartheta_m)$ определяются m функций выборочных значений (условных оценок)

$$\hat{\vartheta}_n^{(i)} = g^{(i)}(x_1, \ldots, x_n), \quad i = 1, \ldots, m, \tag{2.95}$$

в предположении, что значения параметров $\vartheta_1, \ldots, \vartheta_m$ фиксированы, или m усредненных по априорному совместному распределению $w_m(\vartheta_1, \ldots, \vartheta_m)$ параметров функций (безусловных оценок)

$$\hat{\Theta}_n^{(i)} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} g^{(i)}(x_1, \ldots, x_n)\, w_m(\vartheta_1, \ldots, \vartheta_m)\, d\vartheta_1 \ldots d\vartheta_m. \tag{2.95'}$$

Каждая из условных оценок $\hat{\vartheta}_n^{(i)}$ называется состоятельной, если она при $n \to \infty$ сходится по вероятности к параметру ϑ_i, и несмещенной, если для любого n ее среднее по совокупности выборок равно ϑ_i. Аналогично, каждая из безусловных оценок называется состоятельной, если она при $n \to \infty$ сходится по вероятности к $m_1\{\vartheta_i\}$, и несмещенной, если для любого n ее среднее по совокупности выборок равно $m_1\{\vartheta_i\}$.

Оценки $g^{(i)}(x_1, \ldots, x_n)$, $i=1, \ldots, m$, — *совместно достаточные*, если функцию правдоподобия выборки

$$L_x(\boldsymbol{\vartheta}) = W_n(x_1, \ldots, x_n \mid \vartheta_1, \ldots, \vartheta_m) = \prod_{k=1}^{n} w_1(x_k; \vartheta_1, \ldots, \vartheta_m), \tag{2.96}$$
$$\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta),$$

можно представить в виде произведения

$$L_x(\boldsymbol{\vartheta}) = f\,[g^{(1)}(x_1, \ldots, x_n), \ldots$$
$$\ldots, g^{(m)}(x_1, \ldots, x_n), \vartheta_1, \ldots, \vartheta_m]\, h(x_1, \ldots, x_n). \tag{2.97}$$

Пусть $\hat{\vartheta}_n^{(i)} = g^{(i)}(x_1, \ldots, x_n)$ — несмещенная оценка параметра ϑ_i, $i-1, \ldots, m$. Рассмотрим следующие средние по совокупности выборок заданного размера n:

$$I_n^{(i,\,j)}(\vartheta_1, \ldots, \vartheta_m) = m_1\left\{\frac{\partial}{\partial\vartheta_i}\ln L_x(\boldsymbol{\vartheta})\,\frac{\partial}{\partial\vartheta_j}\ln L_x(\boldsymbol{\vartheta})\right\} =$$
$$= -m_1\left\{\frac{\partial^2}{\partial\vartheta_i\,\partial\vartheta_j}\ln L_x(\boldsymbol{\vartheta})\right\}. \tag{2.98}$$

Квадратная матрица размером $m \times m$, элементы которой равны $I_n^{(i,\,j)}$, называется *информационной матрицей Фишера*. Как следует из (2.98), информационная матрица Фишера является корреляционной матрицей совокупности зависимых случайных величин $\frac{\partial}{\partial\vartheta_i}\ln L_x(\boldsymbol{\vartheta})$, $i=1, \ldots, m$, причем в силу (2.37) средние значения этих величин равны нулю.

Если детерминант информационной матрицы отличен от нуля, то имеет место следующее обобщение неравенства Рао — Крамера [см. (2.41)]. Для любых действительных $u_1, \ldots, u_m$ квадратичная форма

$$\sum_{i=1}^{m}\sum_{j=1}^{m}\,[m_1\{(\hat{\vartheta}_n^{(i)}-\vartheta_i)(\hat{\vartheta}_n^{(j)}-\vartheta_j)\} - Y_n^{(i,\,j)}]\,u_i u_j \geqslant 0, \tag{2.99}$$

где $Y_n^{(i,\,j)}$ — элемент матрицы, обратной информационной матрице Фишера (т. е. равен отношению алгебраического дополнения элемента $I_n^{(i,\,j)}$ в матрице Фишера к детерминанту этой матрицы).

Система оценок $\hat{\vartheta}_n^{(1)}, \ldots, \hat{\vartheta}_n^{(m)}$, для которой в (2.99) достигается равенство, называется *совместно эффективной*.

Если это имеет место лишь при $n \to \infty$, то оценки называют совместно асимптотически эффективными.

Элементы информационной матрицы можно выразить через исходное распределение

$$I_n^{(i,j)}(\vartheta, \ldots, \vartheta_m) = n m_1 \left\{ \frac{\partial}{\partial \vartheta_i} \ln w_1(x; \vartheta_1, \ldots, \vartheta_m) \times \right.$$
$$\left. \times \frac{\partial}{\partial \vartheta_j} \ln w_1(x; \vartheta_1, \ldots, \vartheta_m) \right\}. \qquad (2.100)$$

Эта формула аналогична (2.43) и совпадает с ней при $m = 1$.

Используя понятие *корреляционного эллипсоида* (см. приложение IV), можно неравенство (2.99) интерпретировать геометрически: корреляционный эллипсоид любой системы оценок параметров всегда охватывает эллипсоид, уравнение которого

$$\sum_{i=1}^{m} \sum_{j=1}^{m} I_n^{(i,j)} z_i z_j = 1. \qquad (2.101)$$

Корреляционный эллипсоид совместно эффективных оценок совпадает с эллипсоидом (2.101). Отношение квадрата объема эллипсоида (2.101) к квадрату объема корреляционного эллипсоида (см. приложение IV) может быть названо *эффективностью* совместных оценок совокупности параметров.

Рассмотрим теперь методы получения оценок совокупности параметров по определенным критериям качества.

2.4.2. Метод моментов. Этот метод, хотя и не базируется на четко сформулированном критерии качества, используется иногда вследствие своей простоты.

Если исходное распределение $w_1, (x; \vartheta_1, \ldots, \vartheta_s)$ содержит s неизвестных параметров и если существует момент $2s$-го порядка этого распределения, то в силу закона больших чисел все выборочные моменты до s-го порядка включительно, т. е.

$$m_k^* = \frac{1}{n} \sum_{i=1}^{n} x_i^k, \qquad k = 1, \ldots, s,$$

при $n \to \infty$ сходятся по вероятности к соответствующим моментам исходного распределения

$$m_k(\vartheta_1, \ldots, \vartheta_s) = \int_{-\infty}^{\infty} x^k w_1(x; \vartheta_1, \ldots, \vartheta_s)\, dx.$$

Если приравнять выборочные моменты к моментам исходного распределения, то получаем систему s уравнений относительно неизвестных параметров

$$m_k(\vartheta_1, \ldots, \vartheta_s) = \frac{1}{n} \sum_{i=1}^{n} x_i^k, \quad k = 1, \ldots, s, \tag{2.102}$$

решения которой

$$\hat{\vartheta}_n^{(k)} = g^{(k)}(x_1, \ldots, x_n), \quad k = 1, \ldots, s, \tag{2.103}$$

определяют оценки по рассматриваемому методу моментов. При $n \to \infty$ эти оценки асимптотически несмещенны, причем распределение их асимптотически нормальное с дисперсией, убывающей как $\frac{1}{n}$.

Метод моментов может быть использован для любого конечного значения $s \geqslant 1$.

2.4.3. Оценки максимального правдоподобия. Выберем в качестве оценок (условных) неизвестных параметров те значения $\vartheta_1, \ldots, \vartheta_m$, для которых при заданных $x_1, \ldots, x_n$ функция правдоподобия (2.96) как функция m переменных ϑ_i достигает максимума. Условие экстремума записывается системой уравнений

$$\frac{\partial L_x(\boldsymbol{\vartheta})}{\partial \vartheta_i} = 0, \quad i = 1, \ldots, m. \tag{2.104}$$

Однако, как и при оценке одного параметра, вместо (2.104) используется следующая система уравнений правдоподобия:

$$\frac{\partial \ln L_x(\boldsymbol{\vartheta})}{\partial \vartheta_i} = \sum_{k=1}^{n} \frac{\partial}{\partial \vartheta_i} \ln w_1(x_k; \vartheta_1, \ldots, \vartheta_m) = 0,$$

$$i = 1, \ldots, m. \tag{2.105}$$

Если для параметров $\vartheta_1, \ldots, \vartheta_m$ существуют несмещенные совместно эффективные оценки $\hat{\vartheta}_{n\,\text{эф}}^{(1)}, \ldots, \hat{\vartheta}_{n\,\text{эф}}^{(m)}$, то система уравнений правдоподобия имеет *единственное* решение, равное этим оценкам.

При некоторых ограничениях на функцию $w_1(x; \vartheta_1, \ldots, \vartheta_m)$ решение системы уравнений правдоподобия приводит к состоятельным *) асимптотически несмещенным

*) Если при $n \to \infty$ с увеличением размера выборки возрастает и число m неизвестных параметров, то оценки максимального правдоподобия могут и не быть состоятельными (см. [2, § 45]).

и совместно эффективным оценкам, совместное распределение которых асимптотически нормальное со средними, равными оцениваемым параметрам, и корреляционной матрицей, обратной информационной матрице Фишера.

2.4.4. Максимум апостериорной плотности вероятности оцениваемых параметров. Пусть неизвестные параметры $\vartheta_1, \ldots, \vartheta_m$ представляют случайные величины, совместная плотность вероятности которых $w_m(\vartheta_1, \ldots, \vartheta_m)$. Условная плотность вероятности этих параметров, когда извлечена выборка $x_1, \ldots, x_n$, равна

$$W_m(\vartheta_1, \ldots, \vartheta_m \mid x_1, \ldots, x_n) = \\ = \frac{w_m(\vartheta_1, \ldots, \vartheta_m) L_x(\boldsymbol{\vartheta})}{\int\limits_{-\infty}^{\infty} \cdots \int\limits_{-\infty}^{\infty} w_m(\vartheta_1, \ldots, \vartheta_m) L_x(\boldsymbol{\vartheta})\, d\vartheta_1 \ldots d\vartheta_m}. \quad (2.106)$$

Выберем в качестве оценок (безусловных) случайных параметров те значения переменных $\vartheta_1, \ldots, \vartheta_m$, для которых при заданных $x_1, \ldots, x_n$ апостериорная плотность вероятности (2.106) максимальна. Эти значения являются решениями следующей системы уравнений [см. (2.68) и (2.69)]:

$$\frac{\partial}{\partial \vartheta_i} \ln [w_m(\vartheta_1, \ldots, \vartheta_m) L_x(\boldsymbol{\vartheta})] = 0 \quad (2.107)$$

или

$$\frac{\partial}{\partial \vartheta_i} \ln w_m(\vartheta_1, \ldots, \vartheta_m) + \\ + \sum_{k=1}^{n} \frac{\partial}{\partial \vartheta_i} \ln w_1(x_k \mid \vartheta_1, \ldots, \vartheta_m) = 0, \quad i = 1, \ldots, m. \quad (2.108)$$

Оценки, соответствующие максимальной апостериорной плотности, обладают свойствами, схожими со свойствами условных оценок максимального правдоподобия, указанными в предыдущем разделе. Если $\hat{\Theta}_n^{(1)}, \ldots, \hat{\Theta}_n^{(m)}$ — несмещенные совместно эффективные безусловные оценки случайных параметров $\vartheta_1, \ldots, \vartheta_m$, то эти оценки являются единственным решением системы уравнений (2.107).

В некоторых случаях при m неизвестных параметрах представляет интерес оценка лишь $s < m$ параметров, остальные $m - s$ параметров желательно вообще исключить из рассмотрения. Пусть, например, неизвестные m

параметров перенумерованы так, что подлежащими оценке являются первые s параметров. Тогда из (2.106) интегрированием по переменным $\vartheta_{s+1}, \ldots, \vartheta_m$ находится апостериорная плотность вероятности этих параметров

$$W_s(\vartheta_1, \ldots, \vartheta_s \mid x_1, \ldots, x_n) =$$
$$= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} W_m(\vartheta_1, \ldots, \vartheta_s, \vartheta_{s+1}, \ldots, \vartheta_m \mid x_1, \ldots, x_n) \times$$
$$\times d\vartheta_{s+1} \ldots d\vartheta_m,$$

и далее для получения совместных оценок $\hat{\vartheta}_n^{(1)}, \ldots, \hat{\vartheta}_n^{(s)}$ максимизируется условная функция $W_s(\vartheta_1, \ldots, \vartheta_s \mid x_1, \ldots, x_n)$

2.4.5. Байесовские оценки. Когда оценивают m неизвестных параметров, появляются m ошибок, каждой из которых приписывают определенный вес. Функция потерь, учитывающая последствия, которые несут за собой ошибки, будет в общем зависеть от $2m$ переменных:

$$\Pi = \Pi(\hat{\vartheta}_n^{(1)}, \ldots, \hat{\vartheta}_n^{(m)}, \vartheta_1, \ldots, \vartheta_m). \tag{2.109}$$

Иногда выбирают в качестве функции потерь сумму функций потерь определенного вида для отдельных параметров, не учитывая при этом, что потери из-за ошибок для одних параметров более весомы, чем для других. Таким образом, могут быть аналогично (2.71) — (2.74) введены: простая функция потерь

$$\Pi = mc - \sum_{k=1}^{m} \delta(\hat{\vartheta}_n^{(k)} - \vartheta_k); \tag{2.110}$$

функция потерь, суммирующая модули ошибок,

$$\Pi = \sum_{k=1}^{m} |\hat{\vartheta}_n^{(k)} - \vartheta_k|; \tag{2.111}$$

квадратичная функция потерь

$$\Pi = \sum_{k=1}^{m} (\hat{\vartheta}_n^{(k)} - \vartheta_k)^2. \tag{2.112}$$

Для прямоугольной функции потерь [см. (2.74)] может быть принят аддитивный или мультипликативный закон комбинирования отдельных составляющих. Соответственно

$$\Pi = \sum_{k=1}^{m} \Pi_k \tag{2.113}$$

или

$$\Pi = \prod_{k=1}^{m} \Pi_k, \tag{2.113'}$$

где

$$\Pi_k = \begin{cases} 1, & |\hat{\vartheta}_n^{(k)} - \vartheta_k| \geqslant \varepsilon, \\ 0, & |\hat{\vartheta}_n^{(k)} - \vartheta_k| < \varepsilon. \end{cases} \tag{2.113''}$$

Условные байесовские оценки неизвестных параметров $\vartheta_1, \ldots, \vartheta_m$ минимизируют условную функцию риска

$$r(\vartheta_1, \ldots, \vartheta_m) = \\ = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \Pi(\hat{\vartheta}_n^{(1)}, \ldots, \hat{\vartheta}_n^{(m)}, \vartheta_1, \ldots, \vartheta_m) L_x(\boldsymbol{\vartheta})\, dx_1 \ldots dx_n. \tag{2.114}$$

Если оцениваемые параметры случайны, то критерием качества служит минимум среднего риска, т. е. минимум величины

$$R = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} w_m(\vartheta_1, \ldots, \vartheta_m)\, r(\vartheta_1, \ldots, \vartheta_m)\, d\vartheta_1 \ldots d\vartheta_m, \tag{2.115}$$

которая может быть также записана в виде

$$R = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} J(\hat{\vartheta}_n^{(1)}, \ldots, \hat{\vartheta}_n^{(m)} \,|\, x_1, \ldots, x_n) \times \\ \times W_n(x_1, \ldots, x_n)\, dx_1 \ldots dx_n, \tag{2.116}$$

где

$$J(\hat{\vartheta}_n^{(1)}, \ldots, \hat{\vartheta}_n^{(m)} \,|\, x_1, \ldots, x_n) = \\ = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \Pi(\hat{\vartheta}_n^{(1)}, \ldots, \hat{\vartheta}_n^{(m)}, \vartheta_1, \ldots, \vartheta_m) \times \\ \times W_m(\vartheta_1, \ldots, \vartheta_m \,|\, x_1, \ldots, x_n)\, d\vartheta_1 \ldots d\vartheta_m, \tag{2.117}$$

$$W_n(x_1, \ldots, x_n) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} w_m(\vartheta_1, \ldots, \vartheta_m) L_x(\boldsymbol{\vartheta})\, d\vartheta_1 \ldots d\vartheta_m. \tag{2.117'}$$

Оценки, удовлетворяющие указанному критерию, называются *безусловными байесовскими оценками случайных параметров* $\vartheta_1, \ldots, \vartheta_m$.

Средний риск R представляет многомерный функционал, зависящий от m функций $\hat{\vartheta}_n^{(i)} = g_i(x_1, \ldots, x_n)$, $i = 1, \ldots, n$. Так как они входят лишь в выражение (2.117), то система уравнений, определяющих экстремальные функции функционала (2.116), может быть записана в виде

$$\frac{\partial J}{\partial g_i} = 0, \quad i = 1, \ldots, m, \tag{2.118}$$

при условии, что производные в (2.118) существуют.

Обобщая (2.82), можно отметить, что условные оценки максимального правдоподобия неизвестных параметров $\vartheta_1, \ldots, \vartheta_m$ являются частным видом условных байесовских оценок при простой функции потерь. Аналогично, обобщая (2.84), укажем, что безусловные оценки, соответствующие максимуму апостериорной плотности оцениваемых параметров, являются частным случаем безусловных байесовских оценок при простой функции потерь.

Для квадратичной функции потерь из (2.112) и (2.118) следует, что байесовские оценки имеют следующий вид:

$$\hat{\Theta}_n^{(i)*} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \vartheta_i W_m(\vartheta_1, \ldots, \vartheta_m \mid x_1, \ldots, x_n)\, d\vartheta_1 \ldots d\vartheta_m. \tag{2.119}$$

Формула (2.119) после интегрирования по всем переменным, кроме ϑ_i, может быть переписана в виде

$$\hat{\Theta}_n^{(i)*} = \int_{-\infty}^{\infty} \vartheta_i W_1(\vartheta_i \mid x_1, \ldots, x_n)\, d\vartheta_i. \tag{2.120}$$

Таким образом, при квадратичной функции потерь байесовскую оценку каждого из совокупности зависимых параметров $\vartheta_1, \ldots, \vartheta_m$ можно находить отдельно, как при оценке одного параметра [см. (2.87)].

Могут быть также распространены на случай оценки совокупности произвольного конечного числа параметров результаты § 2.3, относящиеся к иным функциям потерь (см. [9 § 21.2.3]). Наконец, не вызывает принципиальных затруднений формальное обобщение понятия минимаксных

оценок. Система оценок $\hat{\vartheta}_n^{(1)}, \ldots, \hat{\vartheta}_n^{(m)}$ называется минимаксной, если верхняя граница соответствующей ей условной функции риска $r(\vartheta_1, \ldots, \vartheta_m)$ не превосходит верхних границ этой функции многих переменных при любой другой системе оценок.

2.5. ОЦЕНКИ ПАРАМЕТРОВ НЕКОТОРЫХ ОДНОМЕРНЫХ ФУНКЦИЙ РАСПРЕДЕЛЕНИЙ

2.5.1. Условная оценка параметра экспоненциального распределения. В качестве примера однопараметрического распределения рассмотрим экспоненциальное распределение. Плотность вероятности случайной величины, подчиняющейся этому распределению, имеет вид [см. (1.95)]

$$w_1(x; \vartheta) = \frac{1}{\vartheta} e^{-\frac{x}{\vartheta}},$$

$$x > 0, \ \vartheta > 0. \tag{2.121}$$

Среднее значение такой случайной величины равно ϑ, а ее дисперсия ϑ^2 (см. задачу 2.1 в первой книге).

Функция правдоподобия выборки размером n с независимыми элементами в соответствии с (2.10) равна

$$W_n(x_1, \ldots, x_n \mid \vartheta) = \frac{1}{\vartheta^n} \exp\left(-\frac{1}{\vartheta} \sum_{i=1}^{n} x_i\right). \tag{2.122}$$

Сравнивая (2.122) с (2.28), сразу же приходим к выводу, что для оценки неизвестного параметра ϑ экспоненциального распределения необходимо располагать не каждым выборочным значением в отдельности, а лишь их суммой

$$g(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i, \tag{2.123}$$

которая в этом случае является единственной достаточной статистикой. В силу ограниченности дисперсии экспоненциального распределения выборочное среднее $\frac{1}{n} \sum_{i=1}^{n} x_i$ в соответствии с общим результатом (2.23) представляет состоятельную и несмещенную оценку среднего ϑ экспоненциального распределения.

В качестве иллюстрации положений, высказанных в § 2.2.4, заметим, что достаточная статистика (2.123) подчиняет отношение правдоподобия

$$l(x_1, \ldots, x_n \mid \vartheta) = \left(\frac{\vartheta_0}{\vartheta}\right)^n \exp\left[-\left(\frac{1}{\vartheta} - \frac{1}{\vartheta_0}\right) \sum_{i=1}^{n} x_i\right].$$

Воспользуемся (2.43) и найдем информацию по Фишеру для экспоненциального распределения

$$I_n(\vartheta) = n \int_0^\infty \left(\frac{1}{\vartheta} - \frac{x}{\vartheta^2}\right) \frac{1}{\vartheta} e^{-\frac{x}{\vartheta}} dx =$$

$$= \frac{n}{\vartheta^4} \int_0^\infty (\vartheta - x)^2 \frac{1}{\vartheta} e^{-\frac{x}{\vartheta}} dx = \frac{n}{\vartheta^2}. \qquad (2.124)$$

Кроме того, дисперсия выборочного среднего всегда (при независимых элементах выборки размером n) в n раз меньше дисперсии исходного распределения, т. е.

$$M_2\left\{\frac{1}{n}\sum_{i=1}^{n} x_i\right\} = \frac{\vartheta^2}{n}. \qquad (2.125)$$

Из (2.124) и (2.125) следует:

$$M_2\left\{\frac{1}{n}\sum_{i=1}^{n} x_i\right\} = \frac{1}{I_n(\vartheta)}. \qquad (2.126)$$

Таким образом, выборочное среднее

$$\hat{\vartheta}_n = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (2.127)$$

является несмещенной эффективной оценкой параметра ϑ экспоненциального распределения (2.121).

Заметим, что производная логарифма функции правдоподобия (2.122)

$$\frac{\partial}{\partial\vartheta} \ln W_n(x_1, \ldots, x_n \mid \vartheta) = -\frac{n}{\vartheta} + \frac{1}{\vartheta^2}\sum_{i=1}^{n} x_i = \frac{n}{\vartheta^2}(\hat{\vartheta}_n - \vartheta)$$

(2.128)

в точности совпадает с представлением (2.46) [с учетом (2.47″)].

Легко убедиться теперь, что выборочное среднее (2.127) является единственным решением уравнения правдоподобия

$$\frac{\partial}{\partial \vartheta} \ln W_n (x_1, \ldots, x_n \mid \vartheta) = 0,$$

что соответствует общему результату, отмеченному в § 2.3.1.

Так как случайная величина $\frac{2}{\vartheta} \sum_{k=1}^{n} x_k$ распределена по закону χ^2 с $2n$ степенями свободы (см. § 1.3.3), то этому же закону следует и величина $\frac{2n}{\vartheta} \sum_{k=1}^{n} x_k$. Тогда интервальная оценка [см. (2.50)] параметра экспоненциального распределения может быть определена из соотношения

$$P \left\{ 2n (1 - \varepsilon_2) < \frac{2n \hat{\vartheta}_n}{\vartheta} < 2n (1 + \varepsilon_1) \right\} = \gamma. \quad (2.129)$$

Для определения величин ε_1 и ε_2 используем (2.51), (2.51′), полагая $\gamma_1 = \gamma_2 = \frac{1}{2} - \frac{\gamma}{2}$. Тогда

$$1 + \varepsilon_1 = \frac{1}{2n} \chi^2_{\frac{1}{2} - \frac{\gamma}{2}}, \quad (2.130)$$

$$1 - \varepsilon_2 = \frac{1}{2n} \chi^2_{\frac{1}{2} + \frac{\gamma}{2}}, \quad (2.131)$$

где χ^2_δ — процентное отклонение случайной величины, распределенной по закону χ^2 с $2n$ степенями свободы.

2.5.2. Байесовские оценки параметра экспоненциального распределения. Предположим, что параметр ϑ экспоненциального распределения — случайная величина, распределенная также по экспоненциальному закону с известным параметром ϑ_0:

$$w_1 (\vartheta) = \frac{1}{\vartheta_0} e^{-\frac{\vartheta}{\vartheta_0}}, \quad \vartheta > 0, \ \vartheta_0 > 0. \quad (2.132)$$

Начнем с байесовской оценки параметра для простой функции потерь, которая является также безусловной оценкой, соответствующей максимальной апостериорной плотности оцениваемого параметра. Эта оценка находится

из уравнения (2.69), которое с учетом (2.122) и (2.132) принимает следующий вид:

$$-\frac{1}{\vartheta_0}+\frac{1}{\vartheta^2}\sum_{k=1}^{n} x_k-\frac{n}{\vartheta}=0. \qquad (2.133)$$

Из (2.133) находим искомую оценку:

$$\hat{\vartheta}_n=\frac{n\vartheta_0}{2}\left(\sqrt{1+\frac{4}{n\vartheta_0}\cdot\frac{1}{n}\sum_{k=1}^{n} x_k}-1\right). \qquad (2.134)$$

Когда размер выборки неограниченно увеличивается ($n \to \infty$),

$$\hat{\vartheta}_n \sim \frac{1}{n}\sum_{k=1}^{n} x_k. \qquad (2.134')$$

Таким образом, среднее арифметическое является асимптотически байесовской оценкой. При ограниченных размерах выборки алгоритм (2.134) получения байесовской оценки по выборочным значениям *нелинейный*.

К оценке максимального правдоподобия (2.134′) приходим также в случае, когда $\vartheta_0 \to \infty$, что соответствует приблизительно равномерному распределению параметра ϑ.

Определим теперь байесовскую оценку для квадратичной функции потерь, используя (2.87). Так как в рассматриваемом случае

$$W_1(\vartheta\,|\,x_1, \ldots, x_n)=\frac{\frac{1}{\vartheta_0\vartheta^n}\exp\left\{-\left(\frac{\vartheta}{\vartheta_0}+\frac{1}{\vartheta}\sum_{k=1}^{n} x_k\right)\right\}}{\frac{1}{\vartheta_0}\int\limits_0^{\infty}\frac{1}{\vartheta^n}\exp\left\{-\left(\frac{\vartheta}{\vartheta_0}+\frac{1}{\vartheta}\sum_{k=1}^{n} x_k\right)\right\}d\vartheta},$$

то искомая байесовская оценка, равная условному среднему, имеет вид

$$\hat{\vartheta}_n=\frac{\int\limits_0^{\infty}\frac{1}{\vartheta^{n-1}}\exp\left\{-\left(\frac{\vartheta}{\vartheta_0}+\frac{1}{\vartheta}\sum_{k=1}^{n} x_k\right)\right\}d\vartheta}{\int\limits_0^{\infty}\frac{1}{\vartheta^{n}}\exp\left\{-\left(\frac{\vartheta}{\vartheta_0}+\frac{1}{\vartheta}\sum_{k=1}^{n} x_k\right)\right\}d\vartheta}. \qquad (2.135)$$

Так как

$$\int_0^\infty x^{-n} e^{-\left(x+\frac{a^2}{x}\right)} dx = \frac{2}{a^{n-1}} K_{n-1}(2a), \qquad (2.136)$$

где $K_m(z)$ — функция Бесселя второго рода от мнимого аргумента, то из (2.135) имеем

$$\hat{\vartheta}_n = \vartheta_0 \sqrt{\frac{1}{\vartheta_0}\sum_{k=1}^{n} x_k} \; \frac{K_{n-2}\left(2\sqrt{\frac{1}{\vartheta_0}\sum_{k=1}^{n} x_k}\right)}{K_{n-1}\left(2\sqrt{\frac{1}{\vartheta_0}\sum_{k=1}^{n} x_k}\right)}. \qquad (2.137)$$

Из сравнения (2.137) с (2.134) видно, что байесовские оценки при простой и квадратичной функциях стоимости существенно различны, как и следовало ожидать, ввиду того, что условное распределение (2.132) *несимметричное*.

2.5.3. Условные оценки параметров нормального распределения. В качестве примера совместной оценки нескольких параметров одномерного распределения рассмотрим оценки среднего a и дисперсии $\sigma^2 < \infty$ нормального распределения. Функция правдоподобия выборки размером n с независимыми элементами для указанного распределения равна согласно (2.10)

$$L_x(a, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - a)^2\right] =$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n} x_i^2 - 2a\sum_{i=1}^{n} x_i + na^2\right]\right\}. \qquad (2.138)$$

Сравнивая (2.138) с (2.97), непосредственно приходим к выводу, что $\sum_{i=1}^{n} x_i$ и $\sum_{i=1}^{n} x_i^2$ представляют совместно достаточные статистики параметров нормального распределения. В силу ограниченности дисперсии и общих соотношений (2.23) и (2.25), справедливых для любых распределений,

выборочное среднее

$$\hat{a} = g_1(x_1, \ldots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2.139}$$

и выборочная дисперсия, умноженная на $\frac{n}{n-1}$,

$$\hat{\sigma}^2 = g_2(x_1, \ldots, x_n) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \hat{a})^2 \tag{2.140}$$

представляют соответственно состоятельные и несмещенные оценки среднего и дисперсии нормального распределения.

Определим элементы информационной матрицы Фишера. В соответствии с (2.98), учитывая (2.138), получаем:

$$I_n^{(1,1)}(a, \sigma^2) = m_1 \left\{ \left[\frac{\partial}{\partial a} L_x(a, \sigma^2) \right]^2 \right\} =$$

$$= m_1 \left\{ \frac{1}{\sigma^4} \left[\sum_{i=1}^{n} (x_i - a) \right]^2 \right\} = \frac{n}{\sigma^2},$$

$$I_n^{(1,2)}(a, \sigma^2) = I^{(2,1)}(a, \sigma^2) =$$

$$= m_1 \left\{ \frac{\partial}{\partial a} L_x(a, \sigma^2) \frac{\partial}{\partial \sigma^2} L_x(a, \sigma^2) \right\} =$$

$$= m_1 \left\{ \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - a) \left[\frac{1}{2\sigma^4} \sum_{k=1}^{n} (x_k - a)^2 - \frac{n}{2\sigma^2} \right] \right\} = 0,$$

$$I_n^{(2,2)}(a, \sigma^2) = m_1 \left\{ \left[\frac{\partial}{\partial \sigma^2} L_x(a, \sigma^2) \right]^2 \right\} =$$

$$= m_1 \left\{ \left[\frac{1}{2\sigma^4} \sum_{k=1}^{n} (x_k - a)^2 - \frac{n}{2\sigma^2} \right]^2 \right\} = \frac{n}{2\sigma^4}.$$

Таким образом, информационная матрица имеет вид

$$\mathbf{I} = \begin{pmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}. \tag{2.141}$$

Детерминант этой матрицы равен

$$\det \mathbf{I} = \frac{n^2}{2\sigma^6} > 0,$$

а элементы матрицы, обратной информационной,

$$Y_n^{(1,1)}=\frac{\sigma^2}{n}, \quad Y_n^{(2,2)}=\frac{2\sigma^4}{n},$$

$$Y_n^{(1,2)}=Y_n^{(2,1)}=0.$$

С другой стороны, дисперсии и ковариация оценок (2.139) и (2.140) равны:

$$M_2\{\hat{a}\}=\frac{\sigma^2}{n},$$

$$M_2\{\hat{\sigma^2}\}=m_1\left\{\left[\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\hat{a})^2-\sigma^2\right]^2\right\}=\frac{2\sigma^4}{n-1},$$

$$m_1\{(\hat{a}-a)(\hat{\sigma^2}-\sigma^2)\}=$$

$$=\frac{1}{n-1}m_1\left\{(\hat{a}-a)\left[\sum_{i=1}^{n}(x_i-a)-n(\hat{a}-a)^2\right]\right\}=0.$$

Подставляя полученные выражения в (2.99), находим

$$\frac{2\sigma^4}{n(n-1)}u_2^2>0,$$

т. е. оценки (2.139) и (2.140) не являются *совместно* эффективными. Корреляционный эллипсоид этих оценок, который описывается уравнением

$$\frac{z_1^2}{\sigma^2 n}+\frac{z_2^2}{2\sigma^4/(n-1)}=1,$$

не совпадает, а охватывает информационный эллипсоид [см. (2.101)]

$$\frac{z_1^2}{\sigma^2/n}+\frac{z_2^2}{2\sigma^4/n}=1.$$

Найдем оценки максимального правдоподобия. Подставляя (2.138) в (2.105), получаем систему уравнений правдоподобия

$$\frac{1}{\sigma^2}\sum_{k=1}^{n}(x_k-a)=0, \tag{2.142}$$

$$\frac{1}{2\sigma^4}\sum_{k=1}^{n}(x_k-a)^2-\frac{n}{2\sigma^2}=0. \tag{2.142'}$$

Из первого уравнения следует:

$$\hat{a} = \frac{1}{n} \sum_{k=1}^{n} x_k, \tag{2.143}$$

т. е. оценкой максимального правдоподобия среднего нормальной случайной величины является выборочное среднее. Заметим, что уравнение (2.142), по существу, не зависит от второго параметра σ^2 нормального распределения.

Сравнение $M_2\{\hat{a}\}$ с $I_n^{(1,1)}$ показывает, что выборочное среднее является несмещенной эффективной оценкой среднего нормальной случайной величины.

Из второго уравнения (2.142′), подставляя вместо a величину $\hat{a}$, находим оценку максимального правдоподобия дисперсии:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{k=1}^{n} (x_k - \hat{a})^2 = M_2^*. \tag{2.144}$$

Таким образом, оценкой максимального правдоподобия дисперсии нормальной случайной величины является выборочная дисперсия. Эта оценка состоятельная, смещенная, причем согласно (2.84) смещение равно

$$b_n(\sigma^2) = -\frac{\sigma^2}{n}. \tag{2.145}$$

Дисперсия оценки (2.144)

$$M_2\{M_2^*\} = \frac{2\sigma^4}{n}\left(1 - \frac{1}{n}\right). \tag{2.146}$$

Нижняя граница дисперсии оценок параметра σ^2 нормальной случайной величины в соответствии с неравенством Рао — Крамера равна

$$M_2\{\hat{\sigma}^2\}_{min} = \frac{2\sigma^4\left(1 - \frac{1}{n}\right)^2}{n}. \tag{2.146′}$$

Из (2.146) и (2.146′) следует, что эффективность оценки максимального правдоподобия дисперсии нормальной случайной величины равна

$$l = \frac{M_2\{\hat{\sigma}^2\}_{min}}{M_2\{M_2^*\}} = 1 - \frac{1}{n}. \tag{2.147}$$

Таким образом, оценка (2.144) не эффективная, а только асимптотически эффективная, в соответствии с общими свойствами оценок максимального правдоподобия, отмеченными выше.

Заметим, что при известном априори среднем оценка максимального правдоподобия дисперсии нормальной случайной величины имеет вид [см. (2.142′)]

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{k=1}^{n} (x_k - a)^2. \qquad (2.148)$$

Эта оценка, как отмечалось в § 2.2.3, несмещенная, а дисперсия ее равна

$$M_2 \{\hat{\sigma}^2\} = \frac{2\sigma^4}{n}, \qquad (2.148')$$

т. е. величине, обратной информации $I_n^{(2,2)}$. Таким образом, при известном среднем оценка максимального правдоподобия дисперсии нормальной случайной величины эффективная.

До сих пор рассматривались точечные оценки (условные) параметров нормального распределения. Теперь исследуем интервальные оценки этих параметров.

Начнем с интервальной оценки среднего нормальной случайной величины в предположении, что второй параметр σ^2 известен точно. В качестве точечной оценки среднего принимаем выборочное среднее (2.139). Вводя нормированную ошибку оценки

$$\varepsilon = \frac{a - \hat{a}}{\sqrt{\frac{\sigma^2}{n}}}, \qquad (2.149)$$

определим вероятность того, что абсолютное значение этой ошибки не превосходит заданной величины ε_γ:

$$P\{|\varepsilon| < \varepsilon_\gamma\} = \gamma. \qquad (2.150)$$

Из (2.149) непосредственно следует, что нормированная ошибка оценки среднего представляет нормальную случайную величину с нулевым средним и единичной дисперсией. Поэтому (2.150) можно переписать в виде

$$2F(\varepsilon_\gamma) - 1 = \gamma \qquad (2.151)$$

или

$$\varepsilon_\gamma = x_{\frac{1}{2} - \frac{\gamma}{2}}, \qquad (2.151')$$

где $F(x)$ — интеграл Лапласа, а $x_{\frac{1}{2}-\frac{\gamma}{2}}$ — процентная точка нормального распределения (см. приложение I).

Последние два уравнения содержат связь между длиной доверительного интервала $2\varepsilon_\gamma$ для нормированной ошибки и коэффициентом доверия γ. Первое из них используется

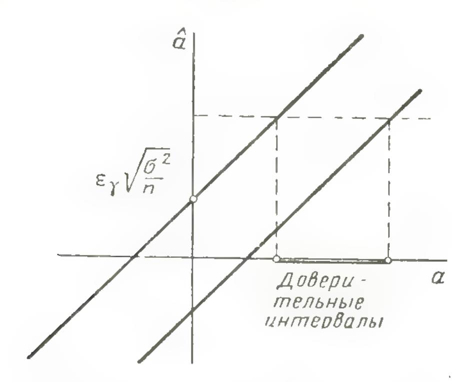


Рис. 16. Построение доверительного интервала.

для определения γ, если задано ε_γ, а второе — для определения ε_γ, когда дано γ. Доверительный интервал для оцениваемого параметра a может быть представлен неравенствами [см. (2.149) и (2.150)]

$$\hat{a} - \varepsilon_\gamma \sqrt{\frac{\sigma^2}{n}} < a < \hat{a} + \varepsilon_\gamma \sqrt{\frac{\sigma^2}{n}}. \qquad (2.152)$$

Связь между $\hat{a}$ и a геометрически представляется двумя прямыми, параллельными биссектрисе координатного угла и отсекающими на оси $\hat{a}$ величины $\pm\varepsilon_\gamma \sqrt{\frac{\sigma^2}{n}}$ (рис. 16). Для определения нижней и верхней границ доверительного интервала необходимо спроектировать на ось абсцисс точки пересечения этих прямых с прямой $\hat{a} = \text{const}$.

Рассмотрим интервальную оценку среднего нормальной случайной величины, когда дисперсия ее неизвестна. В качестве точечных оценок среднего и дисперсии используем несмещенные оценки (2.139) и (2.140). Так же как и в предыдущем случае, введем нормированную ошибку оценки.

Отличие будет состоять в том, что для нормировки используется оценка $\hat{\sigma}^2$, так как точное значение σ^2 неизвестно. Таким образом, в качестве нормированной ошибки оценки среднего выбираем величину *)

$$t = \frac{\hat{a} - a}{\sqrt{\frac{\hat{\sigma}^2}{n}}}, \tag{2.153}$$

которая представляет отношение двух независимых случайных величин: нормальной $(\hat{a} - a)\left(\frac{\sigma^2}{n}\right)^{-1/2}$ (с нулевым средним и единичной дисперсией) и случайной величины $\left(\frac{\hat{\sigma}^2}{\sigma^2}\right)^{1/2}$, распределенной как $\sqrt{\frac{\chi^2}{n-1}}$ (причем χ^2 имеет $(n-1)$ степень свободы). Распределение этого отношения совпадает с *распределением Стьюдента с* $(n-1)$ степенью свободы. Поэтому плотность вероятности нормированной ошибки записывается в виде

$$s_{n-1}(x) = \frac{1}{\sqrt{\pi(n-1)}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \left(1 + \frac{x^2}{n-1}\right)^{-\frac{n}{2}}. \tag{2.154}$$

При увеличении размера выборки n первые два сомножителя в (2.154) приближаются к $(2\pi)^{-1/2}$, а последний — к $e^{-x^2/2}$, так что функция $s_n(x)$ асимптотически стремится при $n \to \infty$ к плотности нормированного отклонения нормальной случайной величины

$$s_n(x) \sim \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{2.154'}$$

Таким образом, для больших размеров выборки можно и при неизвестной дисперсии использовать формулы (2.151), (2.151′), что и следовало ожидать, если иметь в виду состоятельность оценки $\hat{\sigma}^2$. Однако для небольших n распределение Стьюдента заметно отличается от нормального (рис. 17).

Учитывая симметрию распределения Стьюдента, запишем вероятность того, что абсолютное значение нормиро-

---

*) Обозначение t введено по существующей в статистике традиции.

ванной ошибки не больше заданной величины t_γ, в виде

$$P\{|t| < t_\gamma\} = 2 \int_0^{t_\gamma} s_{n-1}(x)\,dx = \gamma. \qquad (2.155)$$

В приложении III приведена таблица значений t_γ, удовлетворяющих соотношению (2.155) при значениях n и γ, изменяющихся в широких пределах. Последняя строка в таблицах соответствует нормальному распределению

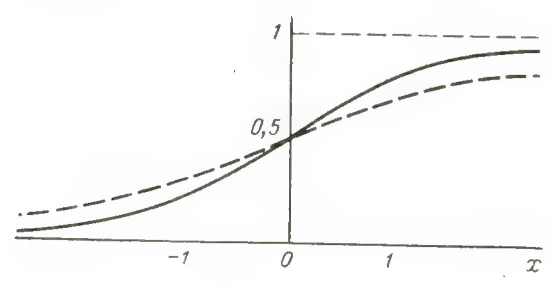


Рис. 17. Распределения:
——— нормальное; — — — Стьюдента при $n = 2$.

($n = \infty$), и значение t_γ (а следовательно, и длина доверительного интервала) в этой строке наименьшее для каждого фиксированного γ. Более широкий доверительный интервал, который получается при одинаковых размерах выборки и коэффициенте доверия по сравнению с предыдущим случаем (известной дисперсии), является платой за неполную информацию о величине дисперсии σ^2 при оценке среднего a.

Доверительный интервал для оцениваемого параметра представляется теперь неравенствами [см. (2.153) и (2.155)]

$$\hat{a} - t_\gamma \sqrt{\frac{\hat{\sigma}^2}{n}} < a < \hat{a} + t_\gamma \sqrt{\frac{\hat{\sigma}^2}{n}}. \qquad (2.156)$$

В заключение рассмотрим вопрос о построении доверительного интервала для неизвестной дисперсии σ^2 нормальной случайной величины. В качестве точечной оценки дисперсии примем несмещенную оценку (2.140). Случайная величина

$$\frac{n-1}{\sigma^2}\hat{\sigma}^2 = \sum_{k=1}^{n} \left(\frac{x_i - \hat{a}}{\sigma}\right)^2 \qquad (2.157)$$

распределена по закону χ^2 с $n-1$ степенями свободы. Выберем в качестве нижней и верхней доверительных границ величины $\varepsilon_1\hat{\sigma}^2$ и $\varepsilon_2\hat{\sigma}^2$. Вероятность того, что

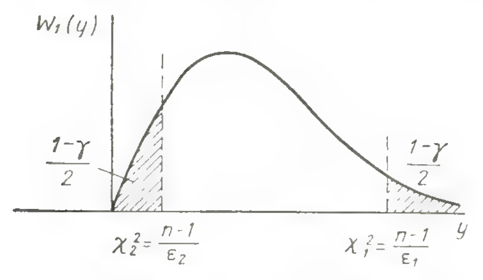


Рис. 18. Определение доверительных границ для дисперсии нормальной случайной величины.

доверительный интервал с указанными границами накроет параметр σ^2, равна

$$P\{\varepsilon_1\hat{\sigma}^2<\sigma^2<\varepsilon_2\hat{\sigma}^2\}=$$
$$=P\left\{\frac{n-1}{\varepsilon_2}<\frac{n-1}{\sigma^2}\hat{\sigma}^2<\frac{n-1}{\varepsilon_1}\right\}=P\left\{\frac{n-1}{\varepsilon_2}<\chi^2<\frac{n-1}{\varepsilon_1}\right\}=\gamma. \tag{2.158}$$

Выберем величины $\varepsilon_1=\frac{n-1}{\chi_1^2}$ и $\varepsilon_2=\frac{n-1}{\chi_2^2}$ из условия (рис. 18)

$$P\{\chi^2<\chi_2^2\}=P\{\chi^2>\chi_1^2\}=\frac{1-\gamma}{2}, \tag{2.159}$$

что равносильно

$$\chi_1^2=\chi^2_{\frac{1}{2}-\frac{\gamma}{2}},\quad \chi_2^2=\chi^2_{\frac{1}{2}+\frac{\gamma}{2}}. \tag{2.159'}$$

Таким образом, доверительный интервал для дисперсии нормальной случайной величины, соответствующий коэффициенту доверия γ, определяется неравенствами

$$\frac{n-1}{\chi^2_{\frac{1}{2}-\frac{\gamma}{2}}}\hat{\sigma}^2<\sigma^2<\frac{n-1}{\chi^2_{\frac{1}{2}+\frac{\gamma}{2}}}\hat{\sigma}^2. \tag{2.160}$$

В приложении II дана таблица процентных точек χ^2_δ *хи*-квадрат распределения.

2.5.4. Метод наименьших квадратов. Пусть производятся измерения некоторой физической величины x, зависящей от некоторой другой величины z, и пусть функциональная зависимость $x = f(z; a_1, a_2, \ldots)$ известна, за исключением некоторых параметров $a_1, a_2, \ldots$ Предположим, что результаты измерений $x_1, \ldots, x_n$ представляют независимые *нормальные* случайные величины, средние значения которых равны $f(z_1; a_1, a_2, \ldots), \ldots f(z_2; a_1, a_2, \ldots)$, а дисперсии одинаковы и равны σ^2 (равноточные, несмещенные измерения). Необходимо воспользоваться результатами измерений для *полного* определения функциональной зависимости между физическими величинами x и z. Совместная функция распределения результатов измерения (функция правдоподобия) равна

$$L_x(a_1, a_2, \ldots) = \frac{1}{\sigma n (2\pi)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2} \times \right.$$

$$\left. \times \sum_{k=1}^{n} [x_k - f(z_k; a_1, a_2, \ldots)]^2 \right\}. \qquad (2.161)$$

Потребуем, чтобы значения неизвестных параметров $a_1, a_2, \ldots$ были такими, которые для заданных результатов измерений $x_1, \ldots, x_n$ максимизируют функцию (2.161). Это равносильно условию минимума величины

$$\sum_{k=1}^{n} [x_k - f(z_k; a_1, a_2, \ldots)]^2,$$

которое используется в классическом *методе наименьших квадратов*.

Значения параметров $a_1, a_2, \ldots$ находятся из системы уравнений

$$\sum_{k=1}^{n} [x_k - f(z_k; a_1, a_2, \ldots)] \frac{\partial f(z_k; a_1, a_2, \ldots)}{\partial a_j} = 0, \quad j = 1, 2, \ldots \qquad (2.162)$$

Если зависимость между физическими величинами x и z *линейная*, т. е.

$$x = a_1 z + a_2, \qquad (2.163)$$

то

$$\frac{\partial f(z_k; a_1, a_2)}{\partial a_1} = z, \quad \frac{\partial f(z_k; a_1, a_2)}{\partial a_2} = 1, \qquad (2.163')$$

а система уравнений (2.162) перепишется в виде

$$\sum_{k=1}^{n} (x_k - a_1 z_k - a_2) z_k = 0,$$

$$\sum_{k=1}^{n} (x_k - a_1 z_k - a_2) = 0,$$

откуда следует:

$$\hat{a}_2 = m_1^* \{x\} - a_1 \langle z \rangle, \qquad (2.164)$$

$$\hat{a}_1 = \frac{m_1^* \{xz\} - m_1^* \{x\} \langle z \rangle}{\langle z^2 \rangle - \langle z \rangle^2}, \qquad (2.164')$$

где $m_1^* \{x\}$, $m_1^* \{xz\}$ — выборочные средние случайных величин x и xz, а

$$\langle z \rangle = \frac{1}{n} \sum_{i=1}^{n} z_i, \quad \langle z^2 \rangle = \frac{1}{n} \sum_{i=1}^{n} z_i^2.$$

Таким образом, метод наименьших квадратов дает следующую линейную зависимость между физическими величинами x и z по данным измерений:

$$x - m_1^* \{x\} = \frac{m_1^* \{xz\} - m_1^* \{x\} \langle z \rangle}{\langle z^2 \rangle - \langle z \rangle^2} (z - \langle z \rangle). \qquad (2.165)$$

Формула (2.165) представляет уравнение эмпирической прямой средней квадратической регрессии (см. (4) в приложении V).

2.5.5. Байесовские оценки параметров нормального распределения. Начнем с того случая, когда дисперсия нормального распределения σ^2 известна точно, а среднее представляет нормальную случайную величину с параметрами (a_0, σ_0^2), т. е.

$$w_1(a) = \frac{1}{\sigma_0 \sqrt{2\pi}} e^{-\frac{(a-a_0)^2}{2\sigma_0^2}}. \qquad (2.166)$$

Найдем оценку, соответствующую максимуму апостериорной плотности вероятности оцениваемого параметра a, которая согласно (2.84) является частным случаем байесовской оценки при простой функции потерь. Используя (2.69), получим уравнение для определения оценки

$$-\frac{a - a_0}{\sigma_0^2} + \frac{1}{\sigma^2} \sum_{k=1}^{n} (x_k - a) = 0,$$

откуда

$$\hat{a}=\frac{1}{1+\frac{\sigma^2}{n\sigma_0}}\left(\frac{1}{n}\sum_{k=1}^{n}x_k+\frac{a_0\sigma^2}{n\sigma_0^2}\right). \qquad (2.167)$$

Оценка (2.167) представляет среднее взвешенное двух величин: оценки максимального правдоподобия $\frac{1}{n}\sum_{k=1}^{n}x_k$ и априорного среднего значения a_0 оцениваемого параметра, причем отношение веса, приписываемого второй величине, к весу первой равно $\frac{\sigma^2}{n\sigma_0^2}$, т. е. отношению дисперсии оценки максимального правдоподобия к априорной дисперсии.

Апостериорная плотность вероятности параметра a для заданной выборки равна

$$W_1(a\,|\,x_1,\ \ldots,\ x_n)=\sqrt{\frac{n}{2\pi\sigma^2}}\left(1+\frac{\sigma^2}{n\sigma_0^2}\right)^{1/2}\times$$
$$\times\exp\left\{-\frac{n}{2\sigma^2}\left(1+\frac{\sigma^2}{n\sigma_0^2}\right)\times\right.$$
$$\left.\times\left[a-\frac{1}{1+\frac{\sigma^2}{n\sigma_0^2}}\left(\frac{1}{n}\sum_{k=1}^{n}x_k+\frac{a_0\sigma^2}{n\sigma_0^2}\right)\right]^2\right\}. \qquad (2.168)$$

При $n\to\infty$ плотность вероятности (2.168) стремится к дельта-функции $\delta\left(a-\frac{1}{n}\sum_{k=1}^{n}x_k\right)$. При данном n и $\frac{\sigma_0}{\sigma}\to\infty$ параметры нормальной функции распределения (2.168) приближаются к $\left(\frac{1}{n}\sum_{k=1}^{n}x_k,\ \frac{\sigma^2}{n}\right)$, а при $\frac{\sigma^2}{n}\gg\sigma_0^2$ (2.168) переходит в (2.166).

Нетрудно также убедиться, что правая часть (2.167) в точности совпадает с условным средним оцениваемого параметра, т. е. с величиной $m_1\{a\,|\,x_1,\ \ldots,\ x_n\}$. Отсюда следует, что (2.167) является байесовской оценкой также и при *квадратичной* функции потерь, что соответствует общему результату, указанному в § 2.3.6 (так как распределение (2.166) симметрично относительно своей единственной моды $a=a_0$). Из § 2.3.8 также следует, что

указанная оценка является байесовской при *симметричной* функции потерь.

Заметим, что

$$\hat{a} \sim \frac{1}{n} \sum_{k=1}^{n} x_k, \tag{2.169}$$

когда

$$\sigma_0^2 \gg \frac{\sigma^2}{n}. \tag{2.169'}$$

Условие (2.169′) выполняется, если при фиксированном $\left(\frac{\sigma}{\sigma_0}\right)^2$ неограниченно увеличивается размер выборки n или если при данном n величина $\sigma_0^2 \gg \sigma^2$. Первое означает, что байесовская оценка (2.167) асимптотически при $n \to \infty$ переходит в оценку максимального правдоподобия. Второе условие можно трактовать следующим образом: распределение $w_1(a)$ неизвестного параметра приблизительно равномерное при сопоставлении его с исходным распределением $w_1(x; a)$. По этой причине оценка (2.167) при $\sigma_0^2 \gg \sigma^2$ переходит в оценку максимального правдоподобия, что и отмечалось в общем виде в конце § 2.3.3.

Отметим также, что в противоположном случае, когда дисперсия исходного распределения σ^2 много больше, чем $n\sigma_0^2$, из (2.167) следует асимптотическое соотношение

$$\hat{a} \sim a_0,$$

т. е. выборочные значения никак не влияют на оценку, которая принимается равной среднему значению оцениваемого параметра.

Подсчитаем величину среднего риска при квадратичной функции потерь. Подставляя (2.167) и (2.168) в (2.79), получаем

$$J(\hat{a} \mid x_1, \ldots, x_n) = \sqrt{\frac{n}{2\pi\sigma^2}} \left(1 + \frac{\sigma^2}{n\sigma_0^2}\right)^{1/2} \times$$

$$\times \int_{-\infty}^{\infty} (a - \hat{a})^2 \exp\left\{-\frac{n}{2\sigma^2}\left(1 + \frac{\sigma^2}{n\sigma_0^2}\right)(a - \hat{a})^2\right\} da =$$

$$= \frac{\sigma^2}{n}\left(1 + \frac{\sigma^2}{n\sigma_0^2}\right)^{-1},$$

т. е. рассматриваемый функционал представляет постоянную величину. Далее из (2.78) находим

$$R = \frac{\sigma^2}{n}\left(1 + \frac{\sigma^2}{n\sigma_0^2}\right)^{-1}. \tag{2.170}$$

Максимальное значение среднего риска соответствует случаю, когда $\sigma_0 \to \infty$, т. е. *равномерное распределение* (в указанном выше смысле) параметра a *наименее благоприятное* и

$$R_{max} = \frac{\sigma^2}{n}. \tag{2.171}$$

Для этого распределения находим функцию риска подстановкой (2.169) в (2.76) с учетом независимости x_i и x_k при $i \neq k$:

$$r(a) = \frac{1}{n^2}\sum_{k=1}^{n}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty}(x_k - a)^2 L_x(a)\,dx_1\dots dx_n = \frac{\sigma^2}{n}. \tag{2.172}$$

Из (2.172) следует, что функция риска при этом не зависит от оцениваемого параметра и ее величина совпадает с максимальным значением среднего риска.

Таким образом, выборочное среднее является также и *минимаксной оценкой* среднего нормальной случайной величины при точно известной ее дисперсии и квадратичной функции потерь.

Рассмотрим теперь байесовскую оценку дисперсии σ^2 нормального распределения при условии, что среднее a известно (положим, что $a = 0$). Зададим априорное распределение дисперсии по экспоненциальному закону с параметром σ_0^2:

$$w_1(\sigma^2) = \frac{1}{\sigma_0^2} e^{-\frac{\sigma^2}{\sigma_0^2}}. \tag{2.173}$$

Найдем оценку, соответствующую максимуму апостериорной плотности вероятности оцениваемого параметра σ^2. Используя (2.69), получаем

$$-\frac{n}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n} x_i^2 - \frac{1}{\sigma_0^2} = 0,$$

откуда

$$\hat{\sigma}^2 = n\sigma_0^2\left[\sqrt{1+\frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{x_i}{\sigma_0}\right)^2}-1\right]. \qquad (2.174)$$

При $n\sigma_0^2 \to \infty$ из (2.174) следует:

$$\hat{\sigma}^2 \sim \frac{1}{n}\sum_{i=1}^{n} x_i^2. \qquad (2.175)$$

Таким образом, при неограниченном увеличении размера выборки или при фиксированном n по мере того, как априорное распределение (2.173) приближается к равномерному, оценка дисперсии, соответствующая максимуму апостериорной плотности, асимптотически стремится к выборочной дисперсии, т. е. к оценке максимального правдоподобия.

Байесовская оценка для квадратичной функции потерь равна [см. (2.87)]

$$\hat{\sigma}^2 = m_1\{\sigma^2 \mid x_1, \ldots, x_n\} =$$

$$= \frac{\int\limits_0^\infty z^{-\frac{n}{2}+1}\exp\left\{-\left(\frac{z}{\sigma_0^2}+\frac{1}{2z}\sum\limits_{i=1}^{n} x_i^2\right)\right\}dz}{\int\limits_0^\infty z^{-\frac{n}{2}}\exp\left\{-\left(\frac{z}{\sigma_0^2}+\frac{1}{2z}\sum\limits_{i=1}^{n} x_i^2\right)\right\}dz}. \qquad (2.176)$$

Используя (2.136), приводим (2.176) к виду

$$\hat{\sigma}^2 - \sigma_0^2\sqrt{\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i}{\sigma_0}\right)^2}\times$$

$$\times\frac{K_{\frac{n}{2}-2}\left[\sqrt{2\sum\limits_{i=1}^{n}\left(\frac{x_i}{\sigma_0}\right)^2}\right]}{K_{\frac{n}{2}-1}\left[\sqrt{2\sum\limits_{i=1}^{n}\left(\frac{x_i}{\sigma_0}\right)^2}\right]}. \qquad (2.177)$$

Как и при оценке параметра экспоненциального распределения, отсутствие симметрии апостериорной плотности вероятности дисперсии приводит к тому, что байесовская оценка (2.177) при квадратичной функции потерь отличается от байесовской оценки (2.174), соответствующей простой функции потерь.

2.6. ОЦЕНКА ОДНОМЕРНОЙ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

2.6.1. Критерии согласия. В предыдущих разделах этой главы изучались методы оценок неизвестных параметров распределения случайной величины в предположении, что функциональная форма закона распределения известна. Часто, однако, ситуация такова, что подобное предположение неоправданно: какая-либо априорная информация о виде указанного закона отсутствует. Необходимо, располагая выборочными значениями, сделать статистический вывод относительно вида функции распределения, из которого извлечена выборка.

Один из принципиальных путей решения этой задачи указывает теорема Гливенко (см. § 2.1.1), согласно которой эмпирическая функция распределения $F_1^*(x)$ сходится по вероятности к гипотетической функции $F_1(x)$, если выборка, по которой построена эмпирическая функция, извлечена из $F_1(x)$. Но эта теорема, устанавливая факт асимптотического приближения $F_1^*(x)$ к $F_1(x)$ при $n \to \infty$, не дает возможность для выборки *конечного размера n* установить, с какими вероятностями могут возникать отклонения эмпирического распределения от гипотетического. Для этого следует прежде всего ввести *количественную меру соответствия* гипотетического и эмпирического распределений, или, как говорят, *критерий согласия.* Этот критерий представляет число $\Delta(F_1^*, F_1)$, которое является функционалом от эмпирического F_1^* и гипотетического F_1 распределений.

Сформулированная выше задача решается в следующей последовательности. По выборочным значениям $x_1, \ldots, x_n$ строится по формуле (2.1) эмпирическая функция распределения $F_1^*(x)$. На основании этого эмпирического распределения, а возможно и других соображений, выдвигается гипотеза о том, что выборка извлечена из распре-

деления $F_1(x)$. По принятому заранее критерию согласия необходимо вычислить $\Delta(F_1^*, F)$ и вероятность того, что указанная величина превосходит некоторый порог Δ_0. Для этого необходимо распределение $\Delta(F_1^*, F_1)$. Определить такое распределение в относительно простой форме удается, как правило, лишь при больших размерах выборки (точнее, удается определить лишь асимптотическое распределение при $n \to \infty$). Если распределение величины $\Delta(F_1^*, F_1)$ известно, то, задаваясь вероятностью α (уровнем значимости критерия) того, что $\Delta > \Delta_0$, находим порог Δ_0. При достаточно малом α получаем хорошее правило проверки гипотезы об истинности гипотетического распределения $F_1(x)$: если для данной выборки $\Delta > \Delta_0$, то гипотеза отвергается, а в противоположном случае — принимается. При этом вероятность отвергнуть правильную гипотезу будет равна уровню значимости (см. гл. 1, стр. 27).

Заметим, что распределение величины $\Delta(F_1^*, F_1)$, определяемое при условии, что выборочные значения $x_1, \ldots, x_n$, по которым построена F_1^*, извлечены из F_1, зависит, вообще говоря, от F_1. Было бы желательно иметь такие критерии согласия, распределения которых не зависели бы от вида гипотетической функции $F_1(x)$. Подобные критерии называют *непараметрическими* (см. § 1.5.3). Рассматриваемые ниже частные виды критериев согласия являются асимптотически (при $n \to \infty$) непараметрическими.

Следует также иметь в виду, что выбранный каким-либо образом критерий согласия дает при достаточно малом уровне значимости α хорошее правило, позволяющее *отвергнуть неверную гипотезу*. Частота отклонения истинной гипотезы в длинной последовательности принятия решений будет равна α. При этом, однако, остается неизвестной частота ошибок другого рода — принятия неверной гипотезы. Если для данной выборки $\Delta < \Delta_0$, то хотя принятое правило предписывает утверждать истинность выдвинутой гипотезы, следует к этому утверждению отнестись более осторожно: с точки зрения заданного критерия согласия нет основания считать, что наблюдаемые выборочные значения не согласуются с гипотетическим распределением.

Недостатком критериев согласия является произвольность выбора самого критерия и уровня значимости, а также трудности, связанные с определением их распределений для конечных размеров выборки. Другой путь состоит

в непосредственном построении оценки неизвестной функции распределения по выборочным значениям без выдвижения гипотезы о виде этой функции (сглаживание эмпирической функции распределения).

2.6.2. Критерий *хи*-квадрат. Этот критерий является одним из наиболее распространенных в приложениях критериев согласия.

Разобьем область, где определена гипотетическая функция распределения $F_1(x)$ случайной величины ξ, на конечное число неперекрывающихся интервалов Δ_i, $i = 1, \ldots, l$. Обозначим

$$p_i = P\{\xi \in \Delta_i\}, \qquad \sum_{i=1}^{l} p_i = 1. \tag{2.178}$$

Пусть в выборке $x_1, \ldots, x_n$ число выборочных значений, попадающих в интервал Δ_i, равно ν_i. Ясно, что $\sum\limits_{i=1}^{l} \nu_i = n$.

Примем в качестве критерия согласия величину

$$\Delta = \sum_{i=1}^{l} \frac{n}{p_i}\left(\frac{\nu_i}{n} - p_i\right)^2 = \sum_{i=1}^{l} \frac{1}{np_i}(\nu_i - np_i)^2. \tag{2.179}$$

Как доказал Пирсон (см. [5, § 30.1]), если проверяемая гипотеза об истинности распределения $F_1(x)$ верна, то при $n \to \infty$ распределение критерия (2.179) асимптотически приближается к χ^2-распределению с $l - 1$ степенями свободы, *не зависящему от вида гипотетического распределения* $F_1(x)$.

Пусть χ^2_α — процентное отклонение случайной величины, имеющей χ^2-распределение с $l - 1$ степенями свободы, т. е. $P\{\chi^2 > \chi^2_\alpha\} = \alpha$ (см. приложение II). При α достаточно малом и большом объеме выборки величина Δ, вычисленная согласно (2.179), практически никогда не будет превосходить порог χ^2_α, если только гипотеза о виде распределения, из которого извлечена выборка, верна. Таким образом, выбирается следующее правило проверки гипотезы: гипотеза отклоняется, если $\Delta > \chi^2_\alpha$, и принимается, если $\Delta \leqslant \chi^2_\alpha$. Вероятность отклонить верную гипотезу равна α.

Кроме отмеченных выше общих недостатков критериев согласия рассмотренный критерий χ^2 обладает еще одним

существенным недостатком, связанным с произвольным разбиением области возможных значений случайной величины на интервалы Δ_i, которое не связано со спецификой функции $F_1(x)$ и соответствующим группированием выборочных значений.

2.6.3. Критерий Колмогорова. Согласно этому критерию количественной мерой соответствия служит для заданного размера n выборки максимум по всем значениям x модуля отклонения эмпирического распределения от гипотетического, т. е.

$$\Delta = \max_x |F_1^*(x) - F_1(x)|. \tag{2.180}$$

Как доказал А. Н. Колмогоров (см. [4], § 62]), если проверяемая гипотеза об истинности $F_1(x)$ верна, то при $n \to \infty$ и дополнительном предположении непрерывности $F_1(x)$, функция распределения величины *) $\Delta\sqrt{n}$ асимптотически приближается к

$$P\{\Delta\sqrt{n} \leqslant z\} \sim \mathrm{k}(z) = \sum_{k=-\infty}^{\infty} (-1)^k \mathrm{e}^{-2k^2z^2}, \quad z > 0. \tag{2.181}$$

Функцию $\mathrm{k}(z)$ можно переписать в виде

$$\mathrm{k}(z) = 1 - 2\sum_{k=1}^{\infty} (-1)^{k-1} \mathrm{e}^{-2k^2z^2}, \quad z > 0. \tag{2.182}$$

С другой стороны, записывая $\mathrm{k}(z)$ в виде

$$\mathrm{k}(z) = \sum_{k=-\infty}^{\infty} \mathrm{e}^{-2(2k)^2z^2} - \sum_{k=-\infty}^{\infty} \mathrm{e}^{-2(2k+1)^2z^2}$$

и используя формулу суммирования Пуассона

$$\sum_{k=-\infty}^{\infty} h(k) = \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} h(x)\, \mathrm{e}^{2\pi ikx}\, dx,$$

---

*) Рассматриваемая функция распределения может быть представлена в виде [12]

$$P\{\Delta\sqrt{n} \leqslant z\} = \sum_{k=-\infty}^{\infty} (-1)^k \mathrm{e}^{-2k^2z^2}\left[1 - \frac{2k^2x}{3\sqrt{n}} + O\left(\frac{1}{n}\right)\right].$$

можно эту функцию представить в виде

$$k(z) = \frac{\sqrt{2\pi}}{z} \sum_{k=1}^{\infty} e^{-\frac{\pi^2 (2k-1)^2}{8z^2}}. \qquad (2.183)$$

Формулу (2.182) удобно использовать при $z > 1$, а (2.183) — при $z \ll 1$.

Если α — заданный уровень значимости, то из (2.182) следует:

$$P\{\Delta > \Delta_\alpha\} = P\{\Delta\sqrt{n} > \Delta_\alpha \sqrt{n}\} \sim 1 - k(\Delta_\alpha \sqrt{n}) =$$

$$= 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 n \Delta_\alpha^2} = \alpha \qquad (2.184)$$

или в другом виде, используя (2.183),

$$P\{\Delta > \Delta\alpha\} \sim \frac{\sqrt{2\pi}}{\Delta_\alpha \sqrt{n}} \sum_{k=1}^{\infty} e^{-\frac{\pi^2 (2k-1)^2}{8n\Delta_\alpha^2}} = \alpha. \qquad (2.185)$$

Ряд в (2.184) сходится быстро, и в качестве первого приближения часто можно ограничиться только первым его членсм, т. е.

$$2e^{-2n\Delta_\alpha^2} = \alpha \qquad (2.186)$$

или

$$\Delta_\alpha = \sqrt{\frac{1}{2n} \ln \frac{2}{\alpha}}. \qquad (2.186')$$

Правило проверки гипотезы таково: если для наблюдаемой выборки $\Delta > \Delta_\alpha$, то гипотеза о том, что выборка извлечена из гипотетического распределения, отвергается; если имеет место противоположное неравенство, то гипотеза принимается. Таблица значений функции k (z) приведена в [4]. Как и критерий χ^2, критерий Колмогорова используется, когда размер выборки очень большой. Однако при использовании этого критерия не требуется предварительного разбиения на интервалы и группирования, как это было в случае критерия χ^2.

Заметим, что при заданном уровне значимости α гипотетическая функция распределения полностью заключена в полосе, границы которой задаются уравнениями

$$y = F_1^*(x) + \Delta$$

и

$$y = F_1^*(x) - \Delta.$$

2.6.4. Критерий Мизеса. Согласно этому критерию количественной мерой соответствия служит для заданного размера n выборки среднее значение квадрата отклонения эмпирического распределения от гипотетического

$$\Delta = \int_{-\infty}^{\infty} [F_1^*(x) - F_1(x)]^2 w_1(x)\, dx, \qquad (2.187)$$

где $w_1(x) = F_1'(x)$ — гипотетическая плотность вероятности случайной величины. Подставляя (2.1) в (2.187) и интегрируя, представим величину Δ в виде *)

$$\Delta = \frac{1}{12n^2} + \frac{1}{n}\sum_{k=1}^{n}\left[F_1(x_k) - \frac{2k-1}{2n}\right]^2. \qquad (2.187')$$

Нетрудно показать [3], что

$$m_1\{\Delta\} = \frac{1}{6n}, \quad M_2\{\Delta\} = \frac{4n-3}{180n^3}. \qquad (2.188)$$

Выражение точного распределения величины $n\Delta$ очень сложное, но при $n > 40$ оно близко к некоторому предельному. Как и критерий Колмогорова, критерий Мизеса, в отличие от χ^2, не связан с группированием выборочных данных, и его распределение достаточно быстро при увеличении размера выборки приближается к предельному. Ниже приведено несколько процентных точек этого предельного распределения:

γ	0,5	0,4	0,3	0,2	0,1	0,05
z_γ	0,1184	0,1467	0,1843	0,2412	0,3473	0,4614
γ	0,03	0,02	0,01	0,001		
z_γ	0,5489	0,6198	0,7435	1,1679		

*) Часто в этом случае вместо Δ используют обозначение ω^2 и критерий согласия называют критерием ω^2.

2.6.5. Принадлежность двух выборок одному и тому же распределению. Предположим, что имеются две выборки $(x_1, \ldots, x_n)$ и $(y_1, \ldots, y_m)$ с независимыми элементами, каждая из которых принадлежит некоторому распределению. Представляет интерес проверить гипотезу о том, что обе эти выборки принадлежат *одному и тому же* распределению. Пусть $F^*_{1x}(z)$ и $F^*_{1y}(z)$ — эмпирические функции распределения, построенные по указанным выше выборкам. В качестве критерия согласия примем величину

$$\Delta = \max_z |F^*_{1x}(z) - F^*_{1y}(z)|. \tag{2.189}$$

Как доказал Н. В. Смирнов (см. [4, § 63]), при $n \to \infty$

$$P\left\{\Delta\left(\frac{1}{n}+\frac{1}{m}\right)^{-\frac{1}{2}} > z\right\} \sim 1 - \mathrm{k}(z), \tag{2.190}$$

где $\mathrm{k}(z)$ — функция, определенная согласно (2.182).

Если α — заданный уровень значимости, то аналогично (2.186′) можно в первом приближении подсчитать порог

$$\Delta_\alpha = \sqrt{\frac{1}{n}+\frac{1}{m}} \cdot \sqrt{\frac{1}{2}\ln\frac{2}{\alpha}} \tag{2.190′}$$

и сформулировать правило проверки гипотезы: выборки принадлежат одному и тому же распределению, если $\Delta < \Delta_\alpha$.

Другой критерий, предложенный Вилькоксоном, основан на подсчете числа *инверсий*. Для этого обе выборки располагают в виде одного *вариационного ряда* (см. § 2.1.1)

$$y_1,\ y_2,\ x_1,\ y_3,\ x_2,\ \ldots$$

Если в этой последовательности заданному x_i предшествует s элементов выборки $y_1, \ldots, y_s$, то имеет место s инверсий. Общее число инверсий U равно сумме инверсий, образуемых всеми элементами первой выборки с элементами второй. Правило проверки гипотезы по критерию Вилькоксона состоит в сравнении общего числа инверсий с пороговым числом, определяемым заданным уровнем значимости.

Можно показать (см. [2, стр. 357]), что при $m+n \geqslant 20$ и $m > 3$ с хорошим приближением можно считать распределение общего числа инверсий *нормальным* с параметрами

$$m_1\{U\} = \frac{mn}{2}, \quad M_2\{U\} = \frac{mn}{12}(m+n+1), \tag{2.191}$$

Тогда пороговое значение U_α числа инверсий для заданного уровня значимости α определяется по формуле

$$U_\alpha = \frac{mn}{2} + x_\alpha \sqrt{\frac{mn}{12}(m+n+1)}, \qquad (2.192)$$

где x_α — процентное отклонение нормальной случайной величины. Если вычисленное по заданным двум выборкам значение U превосходит U_α, то гипотеза о том, что эти выборки принадлежат одному и тому же распределению, отклоняется.

Может использоваться и двухпороговый вариант критерия Вилькоксона, для которого критическая область значений U, когда отклоняется указанная выше гипотеза, определяется двумя порогами:

$$U_{\alpha H} = \frac{mn}{12} - x_{\frac{\alpha}{2}} \sqrt{\frac{mn}{12}(m+n+1)}, \qquad (2.193)$$

$$U_{\alpha b} = \frac{mn}{12} + x_{\frac{\alpha}{2}} \sqrt{\frac{mn}{12}(m+n+1)}. \qquad (2.193')$$

2.6.6. Оценка функции распределения. Если $x_1, \ldots$ $\ldots, x_n$ — выборка с независимыми элементами из неизвестного распределения с плотностью $w_1(x)$, то для оценки этой плотности может быть использована следующая функция (рис. 19):

$$\hat{w}_1(x) = \frac{1}{nh(n)} \sum_{i=1}^{n} K\left[\frac{x - x_i}{h(n)}\right], \qquad (2.194)$$

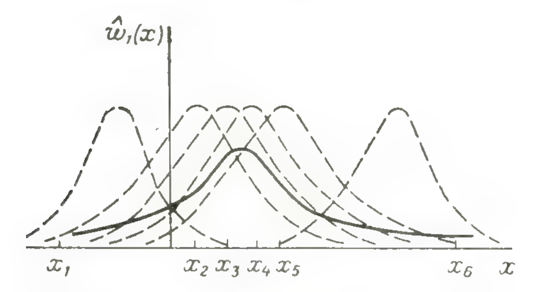


Рис. 19. Построение оценки функции распределения.

где $K(y)$ — произвольное ядро аппроксимации, удовлетворяющее условиям

$$0 < K(y) < \infty, \lim_{y \to \pm\infty} yK(y) = 0, \quad \int_{-\infty}^{\infty} K(y)\,dy = 1, \quad (2.194')$$

и

$$\lim_{n \to \infty} h(n) = 0 \qquad (2.194'')$$

Доказано (см. [15, 17]), что оценка (2.194) состоятельная и *асимптотически несмещенная*. Выбор вида ядра $K(y)$ может быть подчинен некоторым дополнительным условиям оптимальности оценки. В некоторых случаях можно рекомендовать гауссовское ядро

$$K(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.$$

Если вид распределения $w_1(x)$ известен, но неизвестны его параметры, то для оценки функции распределения можно использовать функцию $\hat{w}_1(x)$, параметры которой представляют оценки неизвестных параметров, полученные по выборочным значениям. Например, для оценки нормальной плотности с неизвестными параметрами может быть использована функция

$$\hat{w}_1(x) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} e^{-\frac{(x-\hat{a})^2}{2\hat{\sigma}^2}}, \qquad (2.195)$$

где $\hat{a}$, $\hat{\sigma}^2$ — оценки, определяемые согласно (2.139) и (2.140).

2.7. ОЦЕНКИ ПАРАМЕТРОВ МНОГОМЕРНОЙ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ

2.7.1. Обобщение основных определений на многомерные распределения. Все предыдущее изложение было связано с теорией оценок распределения одной случайной величины. Практический интерес представляют также оценки многомерных распределений или их параметров для совокупности *зависимых* случайных величин $\xi_1, \ldots, \xi_N$. В параметрической теории вид многомерной функции распределения этой совокупности известен и оцениваются лишь

неизвестные параметры, в непараметрической теории оценке подлежит сама функция распределения. В любом случае «сырьем» для статистических выводов является *выборка из многомерного* распределения. В отличие от одномерного случая, элементом такой выборки (размером n) является не одно число, а N чисел

$$x_{1k},\ x_{2k},\ \ldots,\ x_{Nk},\ k = 1,\ \ldots,\ n.$$

Иначе, элемент выборки из многомерного распределения можно представить вектором $\mathbf{x}_k$, имеющим указанные компоненты, а саму выборку — прямоугольной матрицей размером $N \times n$:

$$\mathbf{X} = \| x_{ik} \|.$$

Функция правдоподобия выборки из N-мерного распределения представляет функцию n векторных аргументов, причем параметры этой функции представляют матрицы (в частном случае векторы), элементами которых являются неизвестные параметры многомерного распределения совокупности случайных величин. Для независимых элементов многомерной выборки

$$L_{\mathbf{X}}(\boldsymbol{\vartheta}) = \prod_{k=1}^{n} W_N(x_{1k},\ \ldots,\ x_{Nk};\ \vartheta_1,\ \ldots,\ \vartheta_m), \qquad (2.196)$$

где W_N — функция распределения совокупности случайных величин $\xi_1,\ \ldots,\ \xi_N$.

По выборке $\mathbf{X}$ определяются s выборочных матриц (или векторов), зависящих от выборочных векторов $\mathbf{x}_1,\ \ldots,\ \mathbf{x}_n$,

$$\hat{\mathbf{M}}_i = g(\mathbf{x}_1,\ \ldots,\ \mathbf{x}_n),\ i = 1,\ \ldots,\ s, \qquad (2.197)$$

в предположении, что параметры распределения W_N фиксированы. Эти выборочные матрицы (вектора) представляют *условные оценки* матриц $\mathbf{M}_i$ (векторов), элементами которых являются неизвестные параметры $\vartheta_1,\ \ldots,\ \vartheta_m$.

Если параметры распределения случайны и известно их совместное априорное распределение, то могут быть определены *безусловные оценки*.

Каждая из условных оценок $\hat{\mathbf{M}}_i$ [см. (2.197)] называется состоятельной, если при $n \to \infty$ она сходится по вероятности к $\mathbf{M}_i$. Оценка $\hat{\mathbf{M}}_i$ называется несмещенной, если для любого n ее среднее по совокупности векторных

выборок равно $\mathbf{M}_i$. При этом следует иметь в виду, что под средним значением случайной матрицы (вектора) понимается матрица (вектор), элементы которой равны средним значениям элементов случайной матрицы (вектора). Подобным же образом обобщается и понятие совместно достаточных оценок. Для этого скалярные аргументы в (2.97) заменяются векторными. Информационная матрица Фишера $\mathbf{I}_n$ представляет блочную матрицу, элементами которой являются матрицы

$$\mathbf{I}_n^{(i,\,j)}(\mathbf{M}_1, \ldots, \mathbf{M}_s) = m_1\left\{\frac{\partial}{\partial \mathbf{M}_i} \ln L_{\mathbf{X}}(\mathbf{M}) \frac{\partial}{\partial \mathbf{M}_j} \ln \mathbf{L}_{\mathbf{X}}(\mathbf{M})\right\}, \tag{2.198}$$

$$i,\ j = 1, \ldots, s.$$

Оценки максимального правдоподобия получаются из системы уравнений

$$\frac{\partial \ln \mathbf{L}_{\mathbf{X}}(\mathbf{M})}{\partial \mathbf{M}_i} = 0, \quad i = 1, \ldots, s. \tag{2.199}$$

Аналогично, оценки, соответствующие максимуму апостериорной плотности вероятности оцениваемых параметров, равны

$$\frac{\partial \ln w_s(\mathbf{M}_1, \ldots, \mathbf{M}_s)}{\partial \mathbf{M}_i} + \frac{\partial \ln \mathbf{L}_{\mathbf{X}}(\mathbf{M})}{\partial \mathbf{M}_i} = 0. \tag{2.199'}$$

2.7.2. Оценки вектора средних и корреляционной матрицы многомерного нормального распределения. Прежде чем вводить байесовские оценки параметров многомерного распределения, проиллюстрируем указанные выше обобщения на примере N-мерного нормального распределения. Запишем плотность вероятности этого распределения в векторной форме (см. (2.57) в первой книге):

$$W_N(\mathbf{X}) = \frac{1}{(2\pi)^{\frac{N}{2}}\sqrt{\det \mathbf{M}}} \exp\left[-\frac{1}{2}(\mathbf{x}' - \mathbf{a}')\,\mathbf{M}^{-1}(\mathbf{x} - \mathbf{a})\right], \tag{2.200}$$

где $\mathbf{a}$ — вектор средних значений, $\mathbf{M}$ — корреляционная матрица.

Пусть $\mathbf{x}_1, \ldots, \mathbf{x}_n$ — независимые выборочные N-мерные векторы из приведенного нормального распределения.

Функция правдоподобия (2.196) выборки равна

$$L_{\mathbf{x}}(\mathbf{a}, \mathbf{M}) = \prod_{k=1}^{n} W_N(\mathbf{x}_k; \mathbf{a}, \mathbf{M}) =$$

$$= \frac{1}{(2\pi)^{\frac{nN}{2}} (\det \mathbf{M})^{\frac{n}{2}}} \exp\left[-\frac{1}{2}\sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{a})' \mathbf{M}^{-1} (\mathbf{x}_k - \mathbf{a})\right]. \tag{2.201}$$

Введем вектор выборочных средних

$$\mathbf{m}_1^* = \frac{1}{n}\sum_{k=1}^{n} \mathbf{x}_k \tag{2.202}$$

и выборочную корреляционную матрицу

$$\mathbf{M}^* = \frac{1}{n}\sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{m}_1^*)(\mathbf{x}_k - \mathbf{m}_1^*)'. \tag{2.203}$$

Можно показать, что (см., например, [1, § 3.2])

$$\sum_{k=1}^{n} (\mathbf{x}_k - \mathbf{a})' \mathbf{M}^{-1} (\mathbf{x}_k - \mathbf{a}) =$$

$$= \operatorname{tr}(n\mathbf{M}^{-1}\mathbf{M}^*) + n(\mathbf{m}_1^* - \mathbf{a})' \mathbf{M}^{-1} (\mathbf{m}_1^* - \mathbf{a}), \tag{2.204}$$

где символом tr обозначен след матрицы *).

Подставляя (2.204) в (2.201), представим функцию правдоподобия в виде

$$L_{\mathbf{x}}(\mathbf{a}, \mathbf{M}) = \frac{1}{(2\pi)^{\frac{nN}{2}} (\det \mathbf{M})^{\frac{n}{2}}} \times$$

$$\times \exp\left[-\frac{1}{2} n (\mathbf{m}_1^* - \mathbf{a})' \mathbf{M}^{-1} (\mathbf{m}_1^* - \mathbf{a})\right] \times$$

$$\times \exp\left[-\frac{1}{2}\operatorname{tr}(n\mathbf{M}^{-1}\mathbf{M}^*)\right]. \tag{2.205}$$

Из (2.205) следует, что $\mathbf{m}_1^*$ и $\mathbf{M}^*$ представляют совместно достаточные оценки вектора средних и корреляционной

---

*) Следом квадратной матрицы $\mathbf{A} = \| a_{ij} \|$ называют сумму ее диагональных элементов $\operatorname{tr} A = \sum_i a_{ii}$. Иногда используют другое обозначение вида $\operatorname{sp} A$.

матрицы многомерного нормального распределения. Эти же оценки являются оценками максимального правдоподобия. Оценка $\hat{\mathbf{a}} = \mathbf{m}_1^*$ — несмещенная и эффективная. Оценка $\mathbf{M}^*$ смещенная, так как

$$m_1\{\mathbf{M}^*\} = \left(1 - \frac{1}{n}\right)\mathbf{M}. \tag{2.206}$$

Поэтому несмещенной оценкой корреляционной матрицы $\mathbf{M}$ является величина

$$\hat{\mathbf{M}} = \frac{n}{n-1}\mathbf{M}^* = \frac{1}{n-1}\sum_{k=1}^{n}(\mathbf{x}_k - m_1^*)(\mathbf{x}_k - \mathbf{m}_1^*)'. \tag{2.207}$$

Эффективность оценок $\hat{\mathbf{a}}$ и $\hat{\mathbf{M}}$ (т. е. отношение объемов информационного и корреляционного эллипсоидов) равна $\left(\frac{n-1}{n}\right)\frac{N(N+1)}{2}$.

2.7.3. Байесовские оценки параметров многомерного распределения. Обобщая (2.109), введем функцию потерь для оценки матриц (векторов) многомерного распределения

$$\Pi = \Pi(\hat{\mathbf{M}}_1, \ldots, \hat{\mathbf{M}}_s, \mathbf{M}_1, \ldots, \mathbf{M}_s). \tag{2.208}$$

Условные байесовские оценки минимизируют условную функцию риска

$$r(\mathbf{M}_1, \ldots, \mathbf{M}_s) = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\Pi(\hat{\mathbf{M}}_1, \ldots, \hat{\mathbf{M}}_s, \mathbf{M}_1, \ldots, \mathbf{M}_s)\times$$
$$\times L_{\mathbf{X}}(\mathbf{M})\,dx_1\ldots dx_n, \tag{2.209}$$

Если оцениваемые параметры $\mathbf{M}_1, \ldots, \mathbf{M}_s$ случайны, то, вводя их априорное распределение $w_s(\mathbf{M}_1, \ldots, \mathbf{M}_s)$, запишем величину среднего риска

$$R = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} w_s(\mathbf{M}_1, \ldots, \mathbf{M}_s)\,r(\mathbf{M}_1, \ldots, \mathbf{M}_s)\,d\mathbf{M}_1\ldots d\mathbf{M}_s. \tag{2.210}$$

Безусловные байесовские оценки получаются минимизацией матричного (векторного) функционала (2.210).

2.7.4. Замечание относительно оценки многомерной функции распределения. Критерии согласия для многомерных распределений не исследовались. Естественным образом может быть обобщена на многомерный случай формула (2.194) для оценки неизвестной плотности вероятности $W_N(x_1, \ldots, x_N)$.

Пусть $\mathbf{x}_1, \ldots, \mathbf{x}_n$ — векторы выборочных значений, принадлежащих указанному распределению, и $x_{1k}, \ldots, x_{Nk}$ — компоненты k-го вектора. Если x_k и x_l при $k \neq l$ независимы, то в качестве оценки W_N может быть использована функция

$$\hat{W}_N(x_1, \ldots, x_N) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{N} \frac{1}{h_j(n)} K_j \left[\frac{x_j - x_{ji}}{h_j(n)} \right], \quad (2.211)$$

где $K_j(y)$ — ядро аппроксимации, удовлетворяющее условиям ($j = 1, \ldots, N$)

$$0 < K_j(y) < \infty, \quad \lim_{y \to \pm\infty} y K_j(y) = 0,$$

$$\int_{-\infty}^{\infty} K_j(y)\, dy = 1, \quad \lim_{n \to \infty} h_j(n) = 0. \quad (2.211')$$

Оценка (2.211) состоятельная и асимптотически несмещенная.

Задачи

2.1. Доказать, что выборочные среднее и дисперсия нормальной случайной величины независимы. Обобщить этот результат на многомерное нормальное распределение и доказать, что вектор выборочных средних значений не зависит от выборочной ковариационной матрицы.

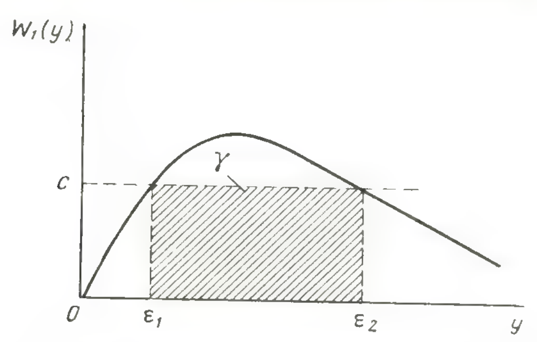


Рис. 20. Определение минимального доверительного интервала.

2.2. Показать, что при заданном коэффициенте доверия γ минимальный доверительный интервал для неизвестной дисперсии нормальной случайной величины определяется из условия (рис. 20)

$$W_1(y) \geqslant c, \quad (1)$$

где $W_1(y)$ — плотность вероятности χ^2 с $n-1$ степенями свободы [см. (2.157)], а константа c определяется из условия

$$\int_{\varepsilon_1}^{\varepsilon_2} W_1(y)\,dy=\gamma, \tag{2}$$

где ε_1 и ε_2 — проекции на ось абсцисс точек пересечения кривой $z=W_1(y)$ с прямой $z=c$ [см. (2.158)].

2.3 [14]. Пусть произведено k групп независимых в совокупности наблюдений, результатом которых являются k выборок размерами $n_1, \ldots, n_k$ из непрерывных распределений $F_{1i}(x)$, $i=1, \ldots, k$. Проверяется гипотеза о том, что $F_{1i}(x)=F_1(x)$ для всех i, и выбирается в качестве критерия согласия величина

$$\Delta=\max_{-\infty<x<\infty}\left[\sum_{i=1}^{k} n_i\{F_{1i}^*(x)-F_0^*(x)\}^2\right]^{\frac{1}{2}}, \tag{3}$$

где

$$F_0^*=\frac{\sum\limits_{i=1}^{k} n_i F_{1i}^*(x)}{\sum\limits_{i=1}^{k} n_i}; \tag{4}$$

$F_{1i}^*(x)$ — эмпирическая функция распределения, построенная по i-й выборке. Показать, что при $n_1, \ldots, n_k\to\infty$

$$P\{\Delta<z\}\sim \mathrm{k}_k(z)=\frac{4}{\Gamma\left(\frac{k-1}{2}\right)(2z^2)^{\frac{k-1}{2}}}\times$$

$$\times\sum_{s=1}^{\infty}\frac{\mu_s^{k-3}}{\left[J'_{\frac{k-3}{2}}(\mu_s)\right]^2}\exp\left[-\frac{\mu_s^2}{2z^2}\right], \qquad z>0, \tag{5}$$

где μ_s — s-й положительный нуль функции Бесселя $J_{\frac{k-3}{2}}(z)$. Убедиться, что при $k=2$

$$\mathrm{k}_2(z)=\mathrm{k}(z)=\frac{\sqrt{2\pi}}{z}\sum_{s=1}^{\infty}\exp\left[-\frac{\pi^2(2s-1)^2}{8z^2}\right], \qquad z>0. \tag{6}$$

2.4. Пусть m — число наступлений события при n независимых испытаниях, в каждом из которых вероятность этого события равна p. Используя неравенство Чебышева, доказать, что для произвольного $\varepsilon>0$

$$\lim_{n\to\infty} P\left\{\left|\frac{m}{n}-p\right|<\varepsilon\right\}=1. \tag{7}$$

2.5 [7]. Необходимо оценить линейную функцию неизвестных параметров a_k, $k=1, \ldots, m$,

$$a=\sum_{k=1}^{m} g_k a_k \tag{8}$$

(g_k — известные величины) по выборочным значениям N случайных величин $x_1, \ldots, x_N$, средние значения которых равны

$$m_1\{x_i\}=\sum_{k=1}^{m} f_{ik} a_k, \quad i=1, \ldots, N, \tag{9}$$

а дисперсии ограничены (известны с точностью до постоянной, т. е. $M_2\{x_i\}=\frac{\sigma^2}{\lambda_i}$, где σ^2 — неизвестно, а λ_i — известно).

Показать, что оценка вида

$$\hat{a}=\sum_{k=1}^{m} g_k \hat{a}_k, \tag{10}$$

где $\hat{a}_k$ — оценка параметра a_k, полученная методом наименьших квадратов (см. § 2.5.4), в классе линейных несмещенных оценок имеет минимальную дисперсию (даже если случайные величины $x_1, \ldots, x_N$ не подчиняются нормальному закону распределения).

2.6. Дисперсия σ^2 нормальной случайной величины известна, а ее среднее значение a распределено по экспоненциальному закону

$$w_1(a)=\frac{1}{a_0} e^{-\frac{a}{a_0}}, \quad a>0, \quad a_0>0. \tag{11}$$

Показать, что байесовская оценка параметра a при простой функции потерь равна

$$\hat{a}_{\text{пр}}=\frac{1}{n}\sum_{k=1}^{n} x_k-\frac{\sigma^2}{n a_0}, \tag{12}$$

а при квадратичной

$$\hat{a}_{\text{кв}}=\hat{a}_{\text{пр}}+\frac{\sigma}{\sqrt{2\pi}}\left[F\left(\frac{\hat{a}_{\text{пр}}}{\sigma}\right)\right]^{-1} e^{-\frac{\hat{a}^2_{\text{пр}}}{2\sigma^2}}, \tag{13}$$

где $F(x)$ — функция Лапласа. Убедиться, что оценка (12) при простой функции потерь соответствует оценке максимальной апостериорной плотности вероятности параметра a и асимптотически при $n\to\infty$ совпадает с оценкой максимального правдоподобия, а оценка (13) при квадратичной функции потерь указанными свойствами не обладает.

ЛИТЕРАТУРА

МОНОГРАФИИ

1. Андерсон Т. Введение в многомерный статистический анализ. Пер. с англ., под ред. Б. В. Гнеденко. Физматгиз, 1963, гл. 2, 3.
2. Ван дер Варден Б. Л. Математическая статистика. Пер. с немецкого, под ред. Н. В. Смирнова. Изд-во иностранной литературы, 1960, гл. VIII.
3. Дунин-Барковский И. В., Смирнов Н. В. Курс теории вероятностей и математической статистики. Изд-во «Наука», 1965, гл. VI, VII.
4. Гнеденко Б. В. Курс теории вероятностей. Физматгиз, 1961, гл. 11.
5. Крамер Г. Математические методы статистики. Пер. с англ., под ред. А. Н. Колмогорова. Изд-во иностранной литературы, 1948, ч. III.
6. Леман Э. Проверка статистических гипотез. Пер. с англ. Ю. В. Прохорова. Изд-во «Наука», 1964, гл. 1.
7. Линник Ю. В. Метод наименьших квадратов и основы теории обработки наблюдений. Физматгиз, 1958, гл. 3, 4.
8. Линник Ю. В. Статистические задачи с мешающими параметрами. Изд-во «Наука», 1966, гл. 2, 3, 7.
9. Миддлтон Д. Введение в статистическую теорию связи. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1962, гл. 21.
10. Хельстром К. Статистическая теория обнаружения сигналов. Пер. с англ., под ред. Ю. Б. Кобзарева. Изд-во иностранной литературы, 1963, гл. 7.

СТАТЬИ

11. Большев Л. Н. Уточнение неравенства Крамера — Рао. «Теория вероятностей и ее применения», 1961, вып. 3.
12. Большев Л. Н. Асимптотически пирсоновские преобразования. «Теория вероятностей и ее применения», 1963, вып. 2.
13. Дынкин Е. Б. Необходимые и достаточные статистики для семейства распределений вероятностей. «Успехи математических наук», 1951, вып. 1.
14. Гихман И. И. Об одном непараметрическом критерии однородности k выборок. «Теория вероятностей и ее применения», 1957, вып. 3.
15. Надарая Э. Я. Некоторые новые оценки функций распределения. «Теория вероятностей и ее применения», 1964, вып. 3.
16. Daniels H. E. The asimptotic efficiency of a maximum likelihood estimator. Proc. 4 Berkl. Symp. on Math. Stat. and Probab., Univ. Calif. Press, 1961. (Русский перевод в журнале «Математика». Изд-во «Мир», 1965, 9 : 1).
17. Parzen E. On estimation of a probability density function and mode. Ann. Math. Statist., 1962, v. 33, № 3.

ГЛАВА ТРЕТЬЯ

СТАТИСТИКА СЛУЧАЙНЫХ ПРОЦЕССОВ

3.1. ДВА СПОСОБА ПРЕДСТАВЛЕНИЯ СЛУЧАЙНОГО ПРОЦЕССА

Существуют два различных подхода к определению случайного процесса (см. § 4.1.2 в первой книге). Если для характеристики вероятностных свойств случайного процесса $\xi(t)$ ограничиваются конечномерной функцией распределения $F_N(x_1, \ldots, x_N, t_1, \ldots, t_N)$, то случайный процесс, по существу, отождествляется с совокупностью N *случайных величин* $\xi_1 = \xi(t_1), \ldots, \xi_N = \xi(t_N)$, которые в общем представляют совокупность *зависимых* случайных величин. Выбор моментов $t_1, \ldots, t_N$ произвольный. Например, если задан интервал наблюдения $(-T, T)$, то можно положить $t_k = \frac{2k-N}{N} T$, $k = 1, \ldots, N$. Результат наблюдения над случайным процессом представляется в рассмотренном случае выборкой из многомерного распределения F_N (см. § 2.7.1).

Другой подход, сохраняющий всю вероятностную информацию о случайном процессе, состоит в рассмотрении совокупности его *реализаций* как функций времени t. Вероятностной мерой, заданной на множестве реализаций, является характеристический функционал случайного процесса [см. (4.9) в первой книге]

$$\Theta[v(t)] = m_1 \left\{ \exp\left[i \int_{-\infty}^{\infty} v(t)\,\xi(t)\,dt \right] \right\},$$

где $v(t)$ — некоторая непрерывная функция и интеграл сходится в среднеквадратическом.

При таком подходе результат наблюдения над случайным процессом представляется конечным числом его усеченных

реализаций

$$\xi_T^{(k)}(t)=\begin{cases}\xi^{(k)}(t), & |t|\leqslant T,\\ 0, & |t|>T,\end{cases}\qquad k=1,\ 2,\ \ldots,\ n.$$

Во многих случаях для получения статистических выводов о характеристиках случайного процесса целесообразно (когда это оказывается возможным) представить случайный процесс в виде суммы *квазидетерминированных* случайных процессов (см. стр. 171 первой книги). Случайные параметры (не зависящие от времени) слагаемых называют иногда *координатами* случайного процесса. Выбор этих координат не однозначен и зависит от конкретных условий задачи.

Рассматривают два вида координат, широко используемых в технических приложениях теории случайных процессов: 1) отсчетные значения процесса с ограниченным энергетическим спектром, 2) некоррелированные координаты.

3.2. ОТСЧЕТ В ДИСКРЕТНЫЕ МОМЕНТЫ ВРЕМЕНИ

3.2.1. Теорема Котельникова (прямая). Рассмотрим детерминированную функцию $f(t)$, спектр (преобразование Фурье) $Z(\omega)$ которой непрерывный и ограничен полосой частот $(-\Delta,\ \Delta)$, т. е. представляет непрерывную *финитную* функцию частоты, равную нулю при $|\omega|\geqslant\Delta$:

$$f(t)=\frac{1}{2\pi}\int\limits_{-\Delta}^{\Delta}Z(\omega)\,e^{i\omega t}\,d\omega. \qquad (3.1)$$

Финитную функцию $Z(\omega)$ можно представить рядом Фурье на интервале $(-\Omega,\ \Omega)$ при условии, что $\Omega\geqslant\Delta$:

$$Z(\omega)=\sum_{n=-\infty}^{\infty}c_n e^{-\frac{2\pi in}{2\Omega}\omega}, \qquad (3.2)$$

где

$$c_n=\frac{1}{2\Omega}\int\limits_{-\Omega}^{\Omega}Z(\omega)\,e^{\frac{2\pi in}{2\Omega}\omega}\,d\omega=$$

$$=\frac{1}{2\Omega}\int\limits_{-\Delta}^{\Delta}Z(\omega)\,e^{\frac{2\pi in}{2\Omega}\omega}\,d\omega=\frac{2\pi}{2\Omega}f\left(\frac{\pi n}{\Omega}\right). \qquad (3.2')$$

Ряд в правой части (3.2) является периодической функцией частоты с периодом 2Ω, которая совпадает с $Z(\omega)$ лишь на основном интервале периодичности $(-\Omega, \Omega)$ и не совпадает на остальных $(\pm k\Omega, \pm (k+1)\Omega)$, $k \geqslant 1$. Подставляя (3.2) в (3.1) и учитывая (3.2′), получим

$$f(t)=\sum_{n=-\infty}^{\infty} f\left(\frac{\pi n}{\Omega}\right) \frac{\sin\left[\Omega\left(t-\frac{\pi n}{\Omega}\right)\right]}{\Omega\left(t-\frac{\pi n}{\Omega}\right)}, \quad \Omega \geqslant \Delta. \qquad (3.3)$$

Выражение (3.3) представляет *интерполяционную* формулу, при помощи которой можно по отсчетам значений функции f в дискретные моменты времени с частотой $2\Omega \geqslant 2\Delta$ восстановить *точно* все значения $f(t)$ на оси времени от $-\infty$ до ∞.

При минимально возможной частоте отбора $2\Omega = 2\Delta$

$$f(t)=\sum_{n=-\infty}^{\infty} f\left(\frac{\pi n}{\Delta}\right) \frac{\sin\left[\Delta\left(t-\frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi n}{\Delta}\right)}. \qquad (3.4)$$

Формула (3.4) составляет содержание известной теоремы В. А. Котельникова, согласно которой детерминированная функция $f(t)$, имеющая ограниченный спектр, полностью определяется своими дискретными значениями в точках, расположенных на расстоянии $\frac{2\pi}{2\Delta}$ друг относительно друга *), где Δ — максимальная частота (круговая) в спектре функции $f(t)$.

Из (3.4) следует, что $f(t)$ получается суперпозицией последовательности функций вида $\frac{\sin \Delta t}{\Delta t}$, каждая из которых получается смещением по оси времени относительно предыдущей на величину интервала отсчета и умножением на отсчетное значение f. Легко видеть, что каждая функция $\sin\left[\Delta\left(t-\frac{\pi n}{\Delta}\right)\right] \Big/ \left[\Delta\left(t-\frac{\pi n}{\Delta}\right)\right]$ в точке отсчета $t=\frac{\pi n}{\Delta}$ имеет максимум, равный единице, и во всех остальных точках отсчета $t=\frac{\pi k}{\Delta}$, $k \neq n$ (k — целое, включая нуль), она равна нулю.

---

*) Заметим что в правой части (3.4) вместо $\frac{\pi n}{\Delta}$ можно подставить $t_0+\frac{\pi n}{\Delta}$, где t_0 — любое фиксированное число (см. задачу 3.4).

Заметим, что из ограниченности спектра функции $f(t)$ следует, что сама функция $f(t)$ не может быть финитной, т. е. нельзя указать интервал значений (даже для очень больших t), на котором функция тождественно равна нулю.

Если функция $f(t)$ задана лишь на конечном интервале времени $(0, T)$, на котором укладывается N отсчетных интервалов, т. е. $\frac{\tilde{N}\pi}{\Delta} \leqslant T < \frac{(N+1)\pi}{\Delta}$, то, доопределив эту функцию вне интервала $(0, T)$ таким образом, чтобы при $n < 0$ и $n > N$ отсчетные значения $f\left(\frac{\pi n}{\Delta}\right) = 0$, получим для этого случая из (3.4)

$$f(t) = \sum_{n=1}^{N} f\left(\frac{\pi n}{\Delta}\right) \frac{\sin\left[\Delta\left(t - \frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t - \frac{\pi n}{\Delta}\right)}, \qquad N = \left[\frac{2\Delta T}{2\pi}\right], \tag{3.5}$$

где [ ] означает наибольшую целую часть числа, заключенного в квадратных скобках. При $N \gg 1$ можно с малой погрешностью полагать $N = \frac{2\Delta T}{2\pi}$.

По отсчетным значениям функции $f(t)$ можно также определить величину интеграла от ее квадрата

$$E = \int_{-\infty}^{\infty} f^2(t)\, dt \tag{3.6}$$

в предположении, что величина E ограничена. Подставляя (3.4) в (3.6) и учитывая, что система функций

$$\frac{\sin\left[\Delta\left(t - \frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t - \frac{\pi n}{\Delta}\right)}, \qquad n = 0, \pm 1, \ldots,$$

ортогональна на интервале $(-\infty, \infty)$ относительно веса $\varphi(t) \equiv 1$, а норма функций равна $\frac{\pi}{\Delta}$ (см. приложение IV

в первой книге), получаем

$$E = \int_{-\infty}^{\infty} f^2(t)\,dt = \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f\left(\frac{\pi n}{\Delta}\right) f\left(\frac{\pi k}{\Delta}\right) \times$$

$$\times \int_{-\infty}^{\infty} \frac{\sin\left[\Delta\left(t-\frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi n}{\Delta}\right)} \frac{\sin\left[\Delta\left(t-\frac{\pi k}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi k}{\Delta}\right)}\,dt =$$

$$= \frac{\pi}{\Delta} \sum_{n=-\infty}^{\infty} f^2\left(\frac{\pi n}{\Delta}\right). \qquad (3.7)$$

Если отсчетные значения $f\left(\frac{\pi n}{\Delta}\right)$ равны нулю при $n < 0$ и $n > N$, то при $N \gg 1$

$$E = \frac{\pi}{\Delta} \sum_{n=0}^{\frac{\Delta T}{\pi}} f^2\left(\frac{\pi n}{\Delta}\right). \qquad (3.7')$$

Отметим также, что формулы (3.5) и (3.7′) можно использовать как приближенные формулы для случая, когда величина

$$1 - \frac{\pi}{E\Delta} \sum_{n=0}^{\frac{\Delta T}{\pi}} f^2\left(\frac{\pi n}{\Delta}\right)$$

пренебрежимо мала.

Наконец, укажем, что при дополнительных ограничениях формула (3.4) верна и в тех случаях, когда спектр функции $f(t)$ состоит из непрерывной части, занимающей ограниченную полосу частот, и дискретной, представляющей сумму дельта-функций (см. § 4.2.5 и 4.2.7 в первой книге):

$$Z(\omega) = Z_{\text{н}}(\omega) + a_0\delta(\omega) + \sum_k a_k\left[\delta(\omega+\omega_k) + \delta(\omega-\omega_k)\right]. \qquad (3.8)$$

Из (3.8) следует, что функция $f(t)$ содержит отличающуюся от нуля постоянную составляющую a_0 и что эта функция содержит гармонические составляющие с периодами $\frac{2\pi}{\omega_k}$. Поэтому формула Котельникова (3.4) для функций со спек-

тром (3.8) должна быть записана в виде

$$f(t)-a_0=\sum_{n=-\infty}^{\infty}\left[f\left(\frac{\pi n}{\Delta}\right)-a_0\right]\frac{\sin\left[\Delta\left(t-\frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi n}{\Delta}\right)}, \quad (3.9)$$

причем $\frac{n}{2\Delta}\neq\frac{1}{\omega_k}$ для любого k, так как в противном случае ряд в правой части (3.9) становится расходящимся при $t=\frac{\pi n}{\Delta}$.

В качестве примера использования теоремы Котельникова приведем интерполяционную формулу для *корреляционной функции* $B(\tau)$ стационарного в широком смысле случайного процесса, энергетический спектр $F(\omega)$ которого непрерывный и тождественно равен нулю при $|\omega|>\Delta$:

$$B(\tau)=\sum_{n=-\infty}^{\infty}B\left(\frac{\pi n}{\Delta}\right)\frac{\sin\left[\Delta\left(\tau-\frac{\pi n}{\Delta}\right)\right]}{\Delta\left(\tau-\frac{\pi n}{\Delta}\right)}. \quad (3.10)$$

(В соответствии с (3.3) в правой части (3.10) вместо Δ можно подставить произвольное $\Omega\geqslant\Delta$.) Используя теорему Хинчина — Винера (см. § 4.2.3 в первой книге) и учитывая, что

$$\int_{-\infty}^{\infty}\frac{\sin\left[\Delta\left(\tau-\frac{\pi n}{\Delta}\right)\right]}{\Delta\left(\tau-\frac{\pi n}{\Delta}\right)}e^{-i\omega\tau}\,d\tau=\begin{cases}\frac{\pi}{\Delta}e^{-\frac{i\pi n\omega}{\Delta}}, & |\omega|\leqslant\Delta,\\ 0, & |\omega|>\Delta,\end{cases} \quad (3.11)$$

получим из (3.10) связь между энергетическим спектром с ограниченной полосой и отсчетными значениями корреляционной функции

$$F(\omega)=\frac{2\pi}{\Delta}\sum_{n=-\infty}^{\infty}B\left(\frac{\pi n}{\Delta}\right)e^{-\frac{i\pi n\omega}{\Delta}}=$$

$$=\frac{2\pi}{\Delta}\left[B(0)+2\sum_{n=1}^{\infty}B\left(\frac{\pi n}{\Delta}\right)\cos\frac{\pi n\omega}{\Delta}\right],\quad |\omega|<\Delta. \quad (3.12)$$

Для белого шума, прошедшего через идеальный фильтр нижних частот (см. стр. 288 в первой книге), корреляцион-

ная функция равна нулю в точках $\tau = \frac{\pi n}{\Delta}$ при $n > 0$. В формулах (3.10) и (3.12) при этом остаются только первые слагаемые, и они переходят в формулы (5.44) и (5.45) первой книги (при $\omega_0 = 0$). Из (3.10) и (3.12) также следует, что стационарный в широком смысле случайный процесс, корреляционная функция которого обращается в нуль при $\tau = \frac{\pi n}{\Delta}$, эквивалентен по энергетическим характеристикам идеально профильтрованному белому шуму.

3.2.2. Теорема Котельникова (обратная). Рассмотрим теперь непрерывную функцию $f(t)$, тождественно равную нулю при $|t| \geqslant \frac{T}{2}$.

Спектр (преобразование Фурье) этой функции равен

$$Z(\omega) = \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t)\, e^{-i\omega t}\, dt. \tag{3.13}$$

Однако непрерывный спектр $Z(\omega)$ функции $f(t)$, ограниченной во времени интервалом $|t| \leqslant \frac{T}{2}$, однозначно определяется заданием спектральных интенсивностей на частотах $\omega_k = \frac{2\pi k}{T}$, где k — любое целое число, включая нуль. Действительно, разлагая на интервале $|t| \leqslant \frac{T}{2}$ финитную функцию $f(t)$ в ряд Фурье, получим *)

$$f(t) = \sum_{n=-\infty}^{\infty} a_n e^{\frac{2\pi i n}{T} t}, \tag{3.14}$$

где

$$a_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t)\, e^{-\frac{2\pi i n}{T} t}\, dt = \frac{1}{T} Z\left(\frac{2\pi i n}{T}\right). \tag{3.14'}$$

---

*) Как и в § 3.2.1, можно получить более общее соотношение, содержащее избыточное число отсчетных значений $Z(\omega)$, разлагая $f(t)$ в ряд Фурье на интервале $|t| < \frac{T_1}{2}$ при условии, что $T_1 \geqslant T$.

Если подставить (3.14) в (3.13) и использовать (3.14′), то после простейших вычислений получим

$$Z(\omega)=\sum_{n=-\infty}^{\infty} Z\left(\frac{2\pi i n}{T}\right)\frac{\sin\left(\frac{\omega T}{2}-\pi n\right)}{\frac{\omega T}{2}-\pi n}. \qquad (3.15)$$

Свойство спектра ограниченной во времени функции, выраженное формулой (3.15), аналогично свойству функции с ограниченным спектром, указываемому теоремой Котельникова [см. (3.4)].

Заметим, что разложение финитной на интервале $\left(-\frac{T}{2}, \frac{T}{2}\right)$ функции $f(t)$ в ряд Фурье (3.14) определяет $f(t)$ через значения ее спектра $Z\left(\frac{2\pi i n}{T}\right)$ на частотах, расположенных на расстоянии $\frac{2\pi}{T}$ друг относительно друга.

Нетрудно найти и аналог формулы (3.7):

$$E=\int_{-\frac{T}{2}}^{\frac{T}{2}} f^2(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}|Z(\omega)|^2\,d\omega=$$

$$=\frac{1}{T}\sum_{n=-\infty}^{\infty}\left|Z\left(\frac{2\pi i n}{T}\right)\right|^2 \qquad (3.16)$$

или с учетом (3.14′)

$$\frac{E}{T}=\sum_{n=-\infty}^{\infty} a_n^2, \qquad (3.16')$$

как и должно быть в соответствии с теоремой Парсеваля (см., например, [8]).

В качестве примера использования обратной теоремы Котельникова приведем интерполяционную формулу для непрерывного энергетического спектра $F(\omega)$ стационарного в широком смысле случайного процесса, корреляционная функция $B(\tau)$ которого тождественно равна нулю при $|\tau|\geqslant\tau_0$:

$$F(\omega)=\sum_{n=-\infty}^{\infty} F\left(\frac{\pi n}{\tau_0}\right)\frac{\sin(\omega\tau_0-\pi n)}{\omega\tau_0-\pi n}. \qquad (3.17)$$

Используя теорему Хинчина — Винера (см. § 4.2.3 в первой книге) и учитывая (3.11), получим из (3.17) связь между финитной корреляционной функцией и отсчетными значениями энергетического спектра:

$$B(\tau) = \frac{1}{4\tau_0} \sum_{n=-\infty}^{\infty} F\left(\frac{\pi n}{\tau_0}\right) e^{\frac{i\pi n\tau}{\tau_0}} =$$

$$= \frac{1}{4\tau_0}\left[F(0) + 2\sum_{n=1}^{\infty} F\left(\frac{\pi n}{\tau_0}\right)\cos\frac{\pi n\tau}{\tau_0}\right], \quad |\tau| < \tau_0. \quad (3.18)$$

3.2.3. Распространение теоремы Котельникова на случайные процессы. Пусть $\xi(t)$ — непрерывный в среднеквадратическом и стационарный в широком смысле случайный процесс, энергетический спектр $F_\xi(\omega)$ которого непрерывный и тождественно равен нулю вне полосы частот $|\omega| < \Delta$. Покажем, что для указанного процесса выполняется равенство (в среднеквадратическом смысле, см. § 3.5 в первой книге)

$$\xi(t) = \sum_{n=-\infty}^{\infty} \xi\left(\frac{\pi n}{\Delta}\right) \frac{\sin\left[\Delta\left(t - \frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t - \frac{\pi n}{\Delta}\right)}. \quad (3.19)$$

Равенство (3.19), обобщающее теорему Котельникова на случайные процессы, означает, что непрерывный в среднеквадратическом смысле процесс с ограниченным энергетическим спектром полностью определяется счетным множеством случайных величин (координат случайного процесса)

$$\xi_n = \xi\left(\frac{\pi n}{\Delta}\right), \quad n = 0, \pm 1, \ldots \quad (3.19')$$

Случайный процесс при указанных ограничениях представляет сумму квазидетерминированных процессов вида

$$\xi_n \frac{\sin\left[\Delta\left(t - \frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t - \frac{\pi n}{\Delta}\right)}.$$

Для доказательства того, что равенство (3.19) справедливо в среднеквадратическом смысле, следует установить равенство корреляционных функций процессов, представ-

ляющих правую и левую части этого равенства. Имеем

$$m_1 \left\{ \sum_{r=-\infty}^{\infty} \xi\left(\frac{\pi r}{\Delta}\right) \frac{\sin\left[\Delta\left(t-\frac{\pi r}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi r}{\Delta}\right)} \times \right.$$

$$\left. \times \sum_{k=-\infty}^{\infty} \xi\left(\frac{\pi k}{\Delta}\right) \frac{\sin\left[\Delta\left(t+\tau-\frac{\pi k}{\Delta}\right)\right]}{\Delta\left(t+\tau-\frac{\pi k}{\Delta}\right)} \right\} =$$

$$= \sum_{r=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} m_1 \left\{ \xi\left(\frac{\pi r}{\Delta}\right) \xi\left(\frac{\pi k}{\Delta}\right) \right\} \frac{\sin\left[\Delta\left(t-\frac{\pi r}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi r}{\Delta}\right)} \times$$

$$\times \frac{\sin\left[\Delta\left(t+\tau-\frac{\pi k}{\Delta}\right)\right]}{\Delta\left(t+\tau-\frac{\pi k}{\Delta}\right)} .$$

В силу стационарности $\xi(t)$ в широком смысле

$$m_1 \left\{ \xi\left(\frac{\pi r}{\Delta}\right) \xi\left(\frac{\pi k}{\Delta}\right) \right\} = B_\xi \left[\frac{\pi}{\Delta}(r-k) \right] .$$

При этом для произвольных t и τ двойная сумма преобразуется в одиночную *)

$$\sum_{r=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} B_\xi \left[\frac{\pi}{\Delta}(r-k) \right] \frac{\sin\left[\Delta\left(t-\frac{\pi r}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi r}{\Delta}\right)} \times$$

$$\times \frac{\sin\left[\Delta\left(t+\tau-\frac{\pi k}{\Delta}\right)\right]}{\Delta\left(t+\tau-\frac{\pi k}{\Delta}\right)} = \sum_{n=-\infty}^{\infty} B_\xi\left(\frac{\pi n}{\Delta}\right) \times$$

$$\times \frac{\sin\left[\Delta\left(\tau-\frac{\pi n}{\Delta}\right)\right]}{\Delta\left(\tau-\frac{\pi n}{\Delta}\right)} . \qquad (3.20)$$

Так как энергетический спектр $F_\xi(\omega)$ ограничен полосой частот $|\omega| \leqslant \Delta$, то в соответствии с (3.10) сумма в правой части (3.20) совпадает с корреляционной функцией $B_\xi(\tau)$.

*) Это непосредственно следует из формулы (6) задачи 3.3.

Используя замечания, приведенные в 3.2.1, можно несколько расширить условия применимости формулы (3.19). Прежде всего эта формула справедлива при замене Δ на любое $\Omega \geqslant \Delta$ [см. (3.3)], а также при замене в правой части $\frac{\pi n}{\Delta}$ на $t_0 + \frac{\pi n}{\Delta}$ (задача 3.4). Кроме того, можно снять ограничение о непрерывности энергетического спектра процесса $\xi(t)$ и предположить, что спектр содержит дискретную часть (сумму дельта-функций) на частотах ω_k, если только $\Delta \neq n\omega_k$. Наконец, возможно обобщение (3.19) на узкополосные случайные процессы (см. задачу 3.1).

Координаты случайного процесса $\xi(t)$, выбираемые согласно теореме Котельникова, являются, вообще говоря, *коррелированными* случайными величинами. Кроме того, использование этих координат ограничено условиями стационарности в широком смысле и финитности энергетического спектра процесса. Переходим к изучению более универсального способа задания координат случайного процесса, свободного от указанных ограничений.

3.3. ОРТОГОНАЛЬНОЕ РАЗЛОЖЕНИЕ СЛУЧАЙНОГО ПРОЦЕССА

3.3.1. Ортогональное разложение корреляционной функции. Из рассмотрения общих свойств корреляционных функций случайных процессов следует, что любая корреляционная функция $B(t, \tau)$ является симметричным ядром положительно полуопределенной квадратичной формы (см. § 4.2 в первой книге), т. е. для любых $n, u_1, \ldots, u_n$,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} B(t_i, t_j)\, u_i u_j \geqslant 0 \tag{3.21}$$

или для любой функции $f(t)$ при условии $\int\limits_{-T}^{T} f^2(t)\, dt < \infty$

$$\int\limits_{-T}^{T} \int\limits_{-T}^{T} B(t, y) f(t) f(y)\, dt\, dy \geqslant 0. \tag{3.21'}$$

Из (3.21) при (3.21′) согласно формуле (9) приложения VI (в первой книге) непосредственно следует возможность представления корреляционной функции случайного про-

цесса в виде следующего ряда:

$$B(t,\ y)=\sum_{k=1}^{\infty}\frac{\varphi_k(t)\,\varphi_k(y)}{\lambda_k}, \tag{3.22}$$

где $\varphi_k(t)$ и λ_k — собственные функции (решения) и собственные числа однородного линейного интегрального уравнения

$$\varphi(t)=\lambda\int_{-T}^{T}B(t,\ y)\,\varphi(y)\,dy, \qquad |t|\leqslant T. \tag{3.23}$$

Собственные функции $\varphi_k(t)$ образуют систему ортогональных нормированных функций, т. е.

$$\int_{-T}^{T}\varphi_k(t)\,\varphi_j(t)\,dt=\begin{cases}0, & k\neq j,\\ 1, & k=j.\end{cases} \tag{3.23'}$$

Умножая обе части (3.23) на $\varphi(t)$, интегрируя затем по t в пределах от $-T$ до T и учитывая (3.21), (3.23'), убеждаемся, что собственные числа $\lambda_k>0$. Кроме того, если корреляционная функция $B(t,\ y)$ положительно определенная (т. е. в (3.21) и (3.21') исключается равенство нулю), то можно показать, что совокупность собственных функций *полная*. Это означает, что на интервале $|t|\leqslant T$ не существует более ни одной функции $\psi(t)$, которая была бы ортогональна всем $\varphi_k(t)$.

Для корреляционной функции $B(\tau)$ стационарного в широком смысле случайного процесса разложение (3.22) может быть переписано в виде

$$B(t-y)=\sum_{k=1}^{\infty}\frac{\varphi_k(t)\,\varphi_k(y)}{\lambda_k}, \tag{3.24}$$

которое при $t=y$ переходит в

$$B(0)=\sum_{k=1}^{\infty}\frac{\varphi_k^2(t)}{\lambda_k}. \tag{3.24'}$$

Из (3.24') интегрированием обеих частей по t с учетом (3.23') получаем

$$\sum_{k=1}^{\infty}\frac{1}{\lambda_k}=2TB(0),$$

т. е. сумма величин, обратных собственным числам, равна средней энергии процесса на интервале $(-T, T)$.

Нетрудно показать, что в общем случае сумма обратных степеней собственных чисел может быть определена по формуле (см. (3.32) в первой книге)

$$\sum_{k=1}^{\infty} \frac{1}{\lambda_k^n} = \int_{-T}^{T} B^{(n)}(u, u)\, du, \qquad (3.24'')$$

где $B^{(n)}(u, v)$ — n-кратная итерация корреляционной функции $B(t, y)$, т. е.

$$B^{(n)}(u, v) = \int_{-T}^{T} \dots \int_{-T}^{T} B(u, t_1) B(t_1, t_2) \dots$$
$$\dots B(t_{n-1}, v)\, dt_1 \dots dt_{n-1}, \quad n \geqslant 2,$$
$$B^{(1)}(u, v) \equiv B(u, v).$$

Заметим, что для корреляционной функции *белого шума* с единичной интенсивностью $B(t, y) = \delta(t - y)$ и из (3.23) следует, что все собственные числа одинаковы, т. е. $\lambda = 1$. Тогда в соответствии с (3.22)

$$\delta(t - y) = \sum_{k=1}^{\infty} \varphi_k(t) \varphi_k(y), \qquad (3.24''')$$

где $\{\varphi_k(t)\}$ — любая система ортонормированных функций.

3.2.2. Некоррелированные координаты случайного процесса. Рассмотрим непрерывный в среднеквадратическом смысле случайный процесс $\xi(t)$ с нулевым средним и корреляционной функцией $B(t, y)$. Выберем в качестве *координат* случайного процесса $\xi(t)$ случайные величины (интеграл в среднеквадратическом)

$$\xi_k = \sqrt{\lambda_k} \int_{-T}^{T} \xi(t) \varphi_k(t)\, dt, \qquad (3.25)$$

где $\varphi_k(t)$ и λ_k — собственные функции и собственные числа интегрального уравнения (3.23). Эти случайные величины имеют, очевидно, нулевые средние. Кроме того, они попарно некоррелированны и имеют одинаковые, равные еди-

нице, дисперсии, так как в силу (3.23) и (3.23′)

$$m_1\{\xi_k\xi_m\} = \sqrt{\lambda_k\lambda_m}\int_{-T}^{T}\int_{-T}^{T} B(t, y)\,\varphi_k(t)\,\varphi_k(y)\,dt\,dy =$$

$$= \sqrt{\frac{\lambda_m}{\lambda_k}}\int_{-T}^{T}\varphi_k(y)\,\varphi_m(y)\,dy = \begin{cases} 0, & k \neq m, \\ 1, & k = m. \end{cases} \tag{3.26}$$

Покажем теперь, что при любом t, принадлежащем интервалу $(-T, T)$, справедливо в среднеквадратическом смысле равенство

$$\xi(t) = \sum_{k=1}^{\infty}\xi_k\frac{\varphi_k(t)}{\sqrt{\lambda_k}}, \tag{3.27}$$

т. е. что

$$\xi_N(t) = \sum_{k=1}^{N}\xi_k\frac{\varphi_k(t)}{\sqrt{\lambda_k}} \tag{3.27′}$$

сходится при $N \to \infty$ в среднеквадратическом к $\xi(t)$.

Находим с учетом (3.25), (3.26) и (3.27′)

$$m_1\{\xi(t)\,\xi_N(t)\} = m_1\left\{\xi(t)\sum_{k=1}^{N}\varphi_k(t)\int_{-T}^{T}\xi(u)\,\varphi_k(u)\,du\right\} =$$

$$= \sum_{k=1}^{N}\int_{-T}^{T}B(t, u)\,\varphi_k(u)\,\varphi_k(t)\,du = \sum_{k=1}^{N}\frac{\varphi_k^2(t)}{\lambda_k},$$

$$m_1\{\xi_N^2(t)\} = \sum_{k=1}^{N}\sum_{n=1}^{N}\frac{\varphi_k(t)\,\varphi_n(t)}{\sqrt{\lambda_k\lambda_n}}\,m_1\{\xi_k\xi_n\} = \sum_{k=1}^{N}\frac{\varphi_k^2(t)}{\lambda_k},$$

откуда следует:

$$m_1\{[\xi(t) - \xi_N(t)]^2\} = B(t, t) - \sum_{k=1}^{N}\frac{\varphi_k^2(t)}{\lambda_k}. \tag{3.28}$$

Из (3.22) при $t = y$ получаем

$$B(t, t) = \sum_{k=1}^{\infty}\frac{\varphi_k^2(t)}{\lambda_k}, \tag{3.29}$$

и, следовательно, при $N \to \infty$ правая часть (3.28) стремится к нулю, т. е. $\xi_N(t)$ сходится в среднеквадратическом к $\xi(t)$.

Таким образом, доказана возможность представления случайного процесса суммой квазидетерминированных случайных процессов вида $\xi_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}$, где $\varphi_k(t)$ и λ_k определяются корреляционной функцией процесса, а ξ_k — случайные величины, получаемые «взвешенным» интегрированием случайного процесса [см. (3.25)].

Если среднее значение $\xi(t)$ отлично от нуля и равно $a(t)$, то разложение (3.27) следует использовать для отклонения случайного процесса от его среднего, и тогда

$$\xi(t) = a(t) + \sum_{k=1}^{\infty} \xi_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}, \quad |t| \leqslant T, \tag{3.30}$$

причем ядром интегрального уравнения (3.23) теперь служит не $B(t, y)$, а $B(t, y) - a(t)a(y)$.

Заметим, что детерминированную функцию $a(t)$ на интервале $|t| \leqslant T$ также можно представить в виде ортогонального разложения

$$a(t) = \sum_{k=1}^{\infty} a_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}, \quad |t| \leqslant T, \tag{3.31}$$

где

$$a_k = \sqrt{\lambda_k} \int_{-T}^{T} a(u) \varphi_k(u)\, du, \tag{3.32}$$

причем $\varphi_k(t)$ и λ_k — *любая* совокупность собственных функций и чисел *). В справедливости (3.31) легко убедиться, подставив в него (3.32) и использовав (3.24‴).

Соединяя (3.30) с (3.31), можно ортогональное разложение случайного процесса $\xi(t)$ с отличным от нуля средним представить в виде

$$\xi(t) = \sum_{k=1}^{\infty} (\xi_k + a_k) \frac{\varphi_k(t)}{\sqrt{\lambda_k}}. \tag{3.33}$$

---

*) Предполагается также, что $\int_{-T}^{T} a^2(t)\, dt = \sum_{k=1}^{\infty} \frac{a_k^2}{\lambda_k} < \infty$.

Отметим, что возможно получить разложение случайного процесса по произвольной совокупности ортогональных детерминированных функций *), но при этом координаты процесса будут, вообще говоря, коррелированы. Только при специальном выборе этих функций, согласованных с корреляционными свойствами процесса [см. (3.23)], координаты становятся некоррелированными. Если к тому же процесс $\xi(t)$ нормальный, то эти некоррелированные координаты становятся *независимыми* случайными величинами.

Основным препятствием к практическому использованию ортогонального разложения случайного процесса (3.27) или (3.30) (т. е. представления его при помощи счетного множества случайных величин) является отыскание точных решений интегрального уравнения (3.23). Лишь для стационарных процессов с дробно-рациональными энергетическими спектрами нахождение этих решений может быть сведено к интегрированию линейного дифференциального уравнения (см., например, приложение 2 в [3]).

3.3.3. Ортогональное разложение комплексного случайного процесса. В некоторых задачах потребуется обобщение указанного ортогонального разложения на комплексный случайный процесс (см. § 4.7.1 в первой книге)

$$\zeta(t) = \xi(t) + i\eta(t).$$

Аналогично (3.22) корреляционную функцию процесса $\zeta(t)$ можно представить в виде

$$B_\zeta(t, y) = m_1\{\zeta(t)\overline{\zeta(y)}\} = \sum_{k=1}^{\infty} \frac{\varphi_k(t)\overline{\varphi_k(y)}}{\lambda_k}, \qquad (3.34)$$

где черта сверху указывает на комплексно-сопряженную величину, а $\varphi_k(t)$ и λ_k — собственные функции и собственные числа интегрального уравнения (3.23) с ядром $B_\zeta(t, y)$.

---

*) Например, гармонических функций $\sqrt{\frac{2}{T}}\cos k\omega_0 t$, $\sqrt{\frac{2}{T}} \times \sin k\omega_0 t$, $\omega_0 = \frac{2\pi}{T}$; в этом случае разложение называют рядом Фурье (в среднеквадратическом).

Тогда, вводя некоррелированные *комплексные* координаты

$$\zeta_k = \sqrt{\lambda_k} \int_{-T}^{T} \zeta(t) \overline{\varphi_k(t)} \, dt, \tag{3.35}$$

$$m_1\{\zeta_k \overline{\zeta_m}\} = \begin{cases} 0, & k \neq m, \\ 1, & k = m, \end{cases} \tag{3.35'}$$

и сохраняя предположение о нулевом среднем, приходим к следующему ортогональному разложению процесса $\zeta(t)$:

$$\zeta(t) = \sum_{k=1}^{\infty} \zeta_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}. \tag{3.36}$$

Если среднее $a_\zeta(t)$ процесса $\zeta(t)$ отлично от нуля, то аналогично (3.33) имеем

$$\zeta(t) = \sum_{k=1}^{\infty} (\zeta_k + a_k) \frac{\varphi_k(t)}{\sqrt{\lambda_k}} \tag{3.37}$$

или

$$a_k = \sqrt{\lambda_k} \int_{-T}^{T} a_\zeta(t) \overline{\varphi_k(t)} \, dt. \tag{3.38}$$

3.3.4. Случай белого шума, прошедшего идеальный фильтр. В качестве примера, иллюстрирующего материал § 3.3.2, рассмотрим стационарный (в широком смысле) случайный процесс на выходе идеального фильтра нижних частот, на вход которого действует белый шум с мощностью N_0 на единицу полосы (см. стр. 288 первой книги). Энергетический спектр этого процесса равен

$$F(\omega) = \begin{cases} 2N_0, & |\omega| \leqslant \Delta, \\ 0, & |\omega| > \Delta, \end{cases} \tag{3.39}$$

а корреляционная функция

$$B(t-y) = B(0) \frac{\sin(t-y)\Delta}{(t-y)\Delta}, \quad B(0) = \frac{N_0 \Delta}{2\pi}. \tag{3.39'}$$

Собственными функциями и собственными числами интегрального уравнения (3.23) с ядром (3.39') являются совокупность ортонормированных функций [см. формулу (6)

задачи 3.3]

$$\varphi_k(t)=\sqrt{\frac{\Delta}{\pi}}\,\frac{\sin\Delta\left(t-\frac{\pi k}{\Delta}\right)}{\Delta\left(t-\frac{\pi k}{\Delta}\right)},\quad -\infty<t<\infty, \tag{3.40}$$

и числа

$$\lambda_k=\frac{1}{N_0}\,. \tag{3.40'}$$

Некоррелированными координатами процесса являются в соответствии с (3.25) случайные величины

$$\xi_k=\sqrt{\frac{\Delta}{\pi N_0}}\int\limits_{-\infty}^{\infty}\xi(t)\,\frac{\sin\left[\Delta\left(t-\frac{\pi k}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi k}{\Delta}\right)}\,dt, \tag{3.41}$$

а его ортогональное разложение, имеющее вид

$$\xi(t)=\sqrt{\frac{\Delta}{\pi N_0}}\sum_{k=-\infty}^{\infty}\xi_k\,\frac{\sin\left[\Delta\left(t-\frac{\pi k}{\Delta}\right)\right]}{\Delta\left(t-\frac{\pi k}{\Delta}\right)}, \tag{3.41'}$$

представляет ряд Котельникова (см. (3.19)), в котором $\xi_k=\sqrt{\frac{\Delta}{\pi N_0}}\,\xi\left(\frac{\pi n}{\Delta}\right)$.

Заметим, что для ортогонального разложения *белого шума* $\xi(t)$ со спектральной интенсивностью N_0 может быть использована любая система ортонормированных функций $\{\varphi_k(t)\}$ (см. (3.24‴)). Некоррелированные координаты белого шума равны

$$\xi_k=\frac{1}{\sqrt{N_0}}\int\limits_{-\infty}^{\infty}\xi(t)\,\varphi_k(t)\,dt. \tag{3.42}$$

3.4. ХАРАКТЕРИСТИКИ НАБЛЮДАЕМЫХ КООРДИНАТ СЛУЧАЙНОГО ПРОЦЕССА

3.4.1. Функция правдоподобия наблюдаемых координат. Предположим, что случайный процесс $\xi(t)$ задается своими координатами, т. е. конечной или счетной совокупностью случайных величин ξ_k. Если задана длительность интервала наблюдения, то результат одного эксперимента может

быть представлен одним элементом выборки из многомерного распределения совокупности $\xi_1, \ldots, \xi_N$. Представление результатов наблюдения может быть различным в зависимости от принятых координат. Так, при наблюдении на интервале $(-T, T)$ через равные промежутки времени $\frac{2T}{N}$ элементом выборки являются N чисел $x_{1k} = \xi^{(1)}(t_k)$, где $t_k = \frac{2k-N}{N} T$, $k = 1, \ldots, N$. Если выбор должен производиться из совокупности некоррелированных случайных величин (3.25), то *наблюдаемыми* координатами будут величины, полученные интегрированием с различными весами $\sqrt{\lambda_k}\varphi_k(t)$ реализации $\xi^{(1)}(t)$ на заданном интервале, т. е.

$$x_{1k} = \sqrt{\lambda_k} \int_{-T}^{T} \xi^{(1)}(t)\, \varphi_k(t)\, dt, \quad k = 1, \ldots, N,$$

где в соответствии с изложенным выше $\varphi_k(t)$ и λ_k определяются корреляционной функцией процесса $\xi(t)$.

При повторных независимых наблюдениях координат случайного процесса получают последовательно элементы выборки $x_{2k}, \ldots, x_{nk}$; $k = 1, \ldots, N$. Совместное распределение выборочных значений — функция правдоподобия наблюдаемых координат — представляет произведение многомерных функций распределения элементов выборки [см. (2.196)]

$$L(x_{11}, \ldots, x_{nN}) = \prod_{k=1}^{n} W_N(x_{k1}, \ldots, x_{kN}). \qquad (3.43)$$

Часто при достаточно большом N удается ограничиться выборкой размером $n = 1$. Заметим, что предел $W_N(x_{k1}, \ldots, x_{kN})$ при $N \to \infty$ *не существует*.

3.4.2. Распределение наблюдаемых координат нормального случайного процесса. Получим выражения функций правдоподобия наблюдаемых координат нормального случайного процесса, которые будут постоянно использоваться во всем дальнейшем изложении.

При наблюдении через равные промежутки времени совместное распределение N наблюдаемых координат, заданное в выборочном пространстве, представляет N-мерную

нормальную функцию распределения (см. (4.159) в первой книге)

$$W_N(x_1, \ldots, x_N) = \frac{1}{(2\pi)^{N/2} \sigma_1 \ldots \sigma_N \sqrt{D}} \times$$

$$\times \exp\left\{-\frac{1}{2D}\sum_{i=1}^{N}\sum_{k=1}^{N} D_{ik}\frac{x_i - a_i}{\sigma_i}\frac{x_k - a_k}{\sigma_k}\right\}, \qquad (3.44)$$

где $a_k = a(t_k)$, $\sigma_k = \sigma(t_k)$ — среднее и среднеквадратическое значения, а D и D_{ik} — детерминант и алгебраические дополнения матрицы коэффициентов корреляции.

Если в качестве координат выбраны случайные величины, определенные согласно (3.25), то выражение функции правдоподобия существенно упрощается. Действительно, если $\xi(t)$ — нормальный случайный процесс с нулевым средним, то его координаты (3.25) представляют нормальные некоррелированные, а следовательно, и независимые случайные величины. Поэтому их многомерное распределение представляет произведение одномерных нормальных функций распределения случайных величин ξ_k, имеющих нулевое среднее и единичную дисперсию, т. е.

$$W_N(x_1, \ldots, x_N) = \frac{1}{(2\pi)^{N/2}} \exp\left\{-\frac{1}{2}\sum_{k=1}^{N} x_k^2\right\}. \qquad (3.45)$$

Для процесса с ненулевым средним $a(t)$ в качестве наблюдаемых координат принимаем в соответствии с (3.33) величину

$$\xi_k + a_k = \sqrt{\lambda_k}\int_{-T}^{T}[\xi(t) + a(t)]\,\varphi_k(t)\,dt, \qquad (3.46)$$

и тогда

$$W_N(x_1, \ldots, x_N) = \frac{1}{(2\pi)^{N/2}} \exp\left\{-\frac{1}{2}\sum_{k=1}^{N}(x_k - a_k)^2\right\}. \qquad (3.47)$$

В том случае, когда имеется n независимых наблюдений реализации на заданном интервале $(-T, T)$ функции правдоподобия могут быть записаны в виде:

для периодического отбора значений нормального случайного процесса

$$L(x_{11}, \ldots, x_{nN}) = \frac{1}{(2\pi)^{nN/2} \sqrt{D^n} (\sigma_1 \ldots \sigma_N)^n} \times$$

$$\times \exp\left\{-\frac{1}{2D} \sum_{r=1}^{n} \sum_{i=1}^{N} \sum_{k=1}^{N} D_{ik} \frac{x_{ri} - a_i}{\sigma_i} \frac{x_{rk} - a_k}{\sigma_k}\right\}, \quad (3.48)$$

для некоррелированных координат

$$L(x_{11}, \ldots, x_{nN}) = \frac{1}{(2\pi)^{nN/2}} \times$$

$$\times \exp\left\{-\frac{1}{2} \sum_{r=1}^{n} \sum_{k=1}^{N} (x_{rk} - a_k)^2\right\}. \quad (3.49)$$

При $n = 1$ формулы (3.48) и (3.49) переходят в (3.44) и (3.47) соответственно.

3.4.3. Отношение правдоподобия и его предельная форма. Как было показано в первой главе, правила выбора решения при проверке статистических гипотез о параметрах распределения случайной величины основаны (для принятых критериев качества) на вычислении отношения правдоподобия и сравнении его с порогом. Очевидно, что при представлении результатов наблюдения над случайным процессом *дискретной выборкой* указанные в первой главе правила решения могут быть *непосредственно* использованы для проверки статистических гипотез относительно параметров распределения случайного процесса. Так, например, если $W_N[x_1, \ldots, x_N \mid s_0(t)]$ и $W_N[x_1, \ldots, x_N \mid s_1(t)]$ — функции правдоподобия выборки ($n = 1$) при гипотезах о том, что среднее равно $s_0(t)$ (гипотеза H_0) и $s_1(t)$ (гипотеза H_1), то правило решения (байесовское, Неймана — Пирсона, иногда и минимаксное) может быть записано в виде: принимается решение H_1, если для наблюдаемых координат $x_1, \ldots, x_N$

$$l(x_1, \ldots, x_N) = \frac{W_N[x_1, \ldots, x_N \mid s_1(t)]}{W_N[x_1, \ldots, x_N \mid s_0(t)]} \geqslant c, \quad (3.50)$$

где порог c зависит от выбранного критерия качества.

Однако при проверке статистических гипотез о параметрах случайных процессов, имеющих непрерывные конечномерные плотности вероятности, можно вместо дискретных

выборок $x_1, \ldots, x_N$ использовать непрерывно наблюдаемую реализацию $x(t)$ случайного процесса $\xi(t)$. Эта возможность базируется на *фундаментальной теореме*, согласно которой при определенных условиях существует конечный предел по вероятности при $N \to \infty$ отношения правдоподобия $l(x_1, \ldots, x_N)$ независимо от того, принадлежит ли совокупность величин $x_1, \ldots, x_N$ распределению $W_N(x_1, \ldots, x_N \mid H_0)$ или $W_N(x_1, \ldots, x_N \mid H_1)$, т. е.

$$\lim_{N\to\infty} l(x_1, \ldots, x_N) = \lim_{N\to\infty} \frac{W_N(x_1, \ldots, x_N \mid H_1)}{W_N(x_1, \ldots, x_N \mid H_0)} = l[x(t)], \tag{3.51}$$

причем предел (3.51) отличен от нуля. Этот предел называется *функционалом отношения правдоподобия* наблюдаемой выборки. Следуя Гренандеру [2], случай, для которого существует функционал отношения правдоподобия $l[x(t)] > 0$, называют *регулярным*.

При других условиях функционал отношения правдоподобия неограниченно возрастает или обращается в нуль, что соответствует вырожденному (или *сингулярному*) случаю.

Можно показать, что в регулярном случае при любом конечном времени наблюдения T правило, использующее функционал отношения правдоподобия, приводит неизбежно к отличным от нуля вероятностям ошибочных решений. Напротив, в сингулярном случае оказываются возможными достоверные безошибочные решения при любом конечном $T > 0$.

Как и в случае дискретной выборки, будет использован не сам функционал отношения правдоподобия, а его логарифм $\ln l[x(t)]$. Тогда в регулярном случае $|\ln l[x(t)]|$ ограничен, а в сингулярном — не ограничен.

Доказательство упомянутой выше теоремы в общем виде и отмеченного следствия этой теоремы требует расширения математического аппарата и выходит за рамки этой книги (отсылаем читателя к [2, 4]). В следующих разделах эти положения будут проиллюстрированы на примере нормального случайного процесса.

3.4.4. Функционал отношения правдоподобия нормального случайного процесса. Пусть $\xi(t)$ — нормальный случайный процесс с нулевым средним и известной корреляционной функцией $B(t, y)$. Этот процесс суммируется

в одном случае с детерминированным процсссом $s_0(t)$, а в другом — с $s_1(t)$. Запишем логарифм отношения правдоподобия для дискретной выборки $x_1, \ldots, x_N$, полученной на интервале $(-T, T)$ через равные промежутки времени из реализации $x(t)$ процесса $\xi(t) + s_0(t)$ или $\xi(t) + s_1(t)$:

$$\ln l(x_1, \ldots, x_N) = -\frac{1}{2D}\sum_{i=1}^{N}\sum_{j=1}^{N} D_{ij}[(x_i - s_{1i})(x_j - s_{1j}) - (x_i - s_{0i})(x_j - s_{0j})] = \frac{1}{2D}\sum_{i=1}^{N}\sum_{j=1}^{N} D_{ij}[(s_{0i} - s_{1i})x_j + (s_{0j} - s_{1j})x_i + (s_{1i}s_{1j} - s_{0i}s_{0j})], \tag{3.52}$$

где

$$x_k = x(t_k); \; s_{0k} = s_0(t_k); \; s_{1k} = s_1(t_k); \; t_k = \frac{2k - N}{N}T;$$

D и D_{ij} — детерминант и алгебраическое дополнение элемента $B(t_i, t_j)$ в корреляционной матрице $\mathbf{M} = \| B(t_i, t_j) \|$, причем предполагается, что матрица $\mathbf{M}$ невырожденная, т. е. $D > 0$, и для любого N существует обратная матрица $\left(\text{элементами которой являются числа } \frac{D_{ij}}{D}\right)$.

Найдем предел (3.52) при $N \to \infty$, т. е. логарифм функционала отношения правдоподобия. Рассмотрим сначала билинейную форму

$$K_N = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{D_{ij}}{D} x_i y_j \tag{3.53}$$

и положим

$$V_j = \sum_{i=1}^{N} \frac{D_{ij}}{D} y_i, \quad j = 1, \ldots, N. \tag{3.54}$$

Тогда

$$K_N = \sum_{j=1}^{N} V_j x_j. \tag{3.54'}$$

Так как $\frac{D_{ij}}{D}$ являются элементами матрицы, обратной корреляционной, то решение системы линейных алгебраиче-

ских уравнений (3.54) может быть представлено в виде

$$y_i = \sum_{j=1}^{N} B(t_i, t_j) V_j, \quad i = 1, \ldots, N. \qquad (3.54'')$$

При $N \to \infty$ (или при $\max(t_{k+1} - t_k) \to 0$ на заданном интервале наблюдения) билинейная форма переходит в интегральную форму (см. приложение VI в первой книге)

$$K = \int_{-T}^{T} V(t)\, x(t)\, dt, \qquad (3.55)$$

а система линейных уравнений (3.54″) — в линейное неоднородное интегральное уравнение

$$y(t) = \int_{-T}^{T} B(t, u)\, V(u)\, du, \ |t| \leqslant T, \qquad (3.56)$$

из которого по заданным $y(t)$ и $B(t, u)$ должно быть найдено неизвестное $V(t)$, входящее в выражение (3.55).

Используя (3.55) и (3.56), можно перейти в (3.52) к пределу при $N \to \infty$ и получить следующее выражение логарифма функционала отношения правдоподобия, к которому по вероятности сходится логарифм отношения правдоподобия $\ln l(x_1, \ldots, x_N)$:

$$\ln l[x(t)] = \int_{-T}^{T} V(t)\, x(t)\, dt + \frac{1}{2} \int_{-T}^{T} [V_0(t)\, s_0(t) - V_1(t)\, s_1(t)]\, dt, \qquad (3.57)$$

где $V(t)$, $V_1(t)$, $V_0(t)$ — решения следующих неоднородных линейных интегральных уравнений (везде $|t| \leqslant T$):

$$\int_{-T}^{T} B(t, u)\, V(u)\, du = s_1(t) - s_0(t), \qquad (3.58)$$

$$\int_{-T}^{T} B(t, u)\, V_1(u)\, du = s_1(t), \qquad (3.59)$$

$$\int_{-T}^{T} B(t, u)\, V_0(u)\, du = s_0(t). \qquad (3.59')$$

Сравнение (3.58) с (3.59) и (3.59') показывает, что

$$V(t) = V_1(t) - V_0(t). \tag{3.60}$$

Далее, умножая обе части (3.59) на $V_0(t)$, а (3.59') — на $V_1(t)$ и интегрируя по t в пределах от $-T$ до T, получим

$$\int_{-T}^{T} s_1(t) V_0(t)\, dt = \int_{-T}^{T} s_0(t) V_1(t)\, dt. \tag{3.60'}$$

Используя (3.60) и (3.60'), нетрудно формулу (3.57) привести к виду

$$\ln l\,[x(t)] = \int_{-T}^{T} V(t) \left[x(t) - \frac{s_0(t) + s_1(t)}{2} \right] dt. \tag{3.61}$$

Выражение (3.61) логарифма функционала отношения правдоподобия можно получить также предельным переходом при $N \to \infty$ в логарифме отношения правдоподобия для N наблюдаемых *независимых* координат. Используя (3.47), получим

$$\ln l(x_1, \ldots, x_N) = -\frac{1}{2} \sum_{k=1}^{N} [(x_k - b_k)^2 - (x_k - a_k)^2] =$$

$$= \frac{1}{2} \sum_{k=1}^{N} (a_k^2 - b_k^2) + \sum_{k=1}^{N} x_k (b_k - a_k), \tag{3.62}$$

где

$$x_k = \sqrt{\lambda_k} \int_{-T}^{T} x(t)\, \varphi_k(t)\, dt; \tag{3.63}$$

$$a_k = \sqrt{\lambda_k} \int_{-T}^{T} s_0(t)\, \varphi_k(t)\, dt; \tag{3.64}$$

$$b_k = \sqrt{\lambda_k} \int_{-T}^{T} s_1(t)\, \varphi_k(t)\, dt; \tag{3.64'}$$

λ_k и $\varphi_k(t)$ — собственные числа и собственные функции однородного линейного интегрального уравнения (3.23).

Обозначим

$$V_N(t)=\sum_{k=1}^{N}\sqrt{\lambda_k}\,(b_k-a_k)\,\varphi_k(t). \qquad (3.65)$$

Тогда выражение (3.62) можно переписать в виде

$$\ln l(x_1,\ \ldots,\ x_N)=\int_{-T}^{T}V_N(t)\left[x(t)-\frac{s_0(t)+s_1(t)}{2}\right]dt. \qquad (3.66)$$

Заметим также, что, умножая обе части (3.65) на $B(t, u)$, интегрируя по t от $-T$ до T и используя (3.23), получаем

$$\int_{-T}^{T}B(t,\ u)\,V_N(t)\,dt=\sum_{k=1}^{N}\frac{b_k-a_k}{\sqrt{\lambda_k}}\,\varphi_k(u). \qquad (3.67)$$

Если

$$\sum_{k=1}^{\infty}\lambda_k\,(a_k-b_k)^2<\infty, \qquad (3.68)$$

то существует конечный предел

$$\lim_{N\to\infty}V_N(t)=V(t), \qquad (3.68')$$

причем функция $V(t)$ определяется из неоднородного линейного интегрального уравнения [см. (3.67) и (3.31)]

$$\int_{-T}^{T}B(t,\ u)\,V(u)\,du=\sum_{k=1}^{\infty}\frac{b_k}{\sqrt{\lambda_k}}\,\varphi_k(t)-$$

$$-\sum_{k=1}^{\infty}\frac{a_k}{\sqrt{\lambda_k}}\,\varphi_k(t)=s_1(t)-s_0(t),\quad |t|\leqslant T,$$

которое не отличается от (3.58).

При выполнении условия (3.68) из (3.68′) следует, что логарифм отношения правдоподобия (3.66) сходится по вероятности к следующему функционалу:

$$\ln l\,[x(t)]=\int_{-T}^{T}V(t)\left[x(t)-\frac{s_0(t)+s_1(t)}{2}\right]dt.$$

Последнее выражение полностью совпадает с (3.61).

Если ряд $\sum_{k=1}^{\infty} \lambda_k (a_k - b_k)^2$ расходящийся, то можно показать (см., например, [2, § 4.4]), что логарифм отношения правдоподобия (3.66) при $N \to \infty$ стремится по вероятности к $+\infty$, если $x(t)$ принадлежит процессу $\xi(t) + s_1(t)$, и стремится к $-\infty$, если $x(t)$ принадлежит процессу $\xi(t) + s_0(t)$.

Таким образом, при выполнении условия (3.68) имеет место регулярный случай (по терминологии, указанной в § 3.4.3), а если это условие не выполняется — то сингулярный.

Если имеется не одна, а n наблюдаемых независимых реализаций $x_1(t), \ldots, x_n(t)$, то, учитывая (3.48) или (3.49), можно непосредственно обобщить (3.61), представив логарифм функционала отношения правдоподобия в виде

$$\ln l\,[x_1(t), \ldots, x_n(t)] = \int_{-T}^{T} V(t) \left\{ \sum_{i=1}^{n} x_i(t) - \frac{n}{2}\,[s_0(t) + s_1(t)] \right\} dt, \qquad (3.69)$$

где $V(t)$ по-прежнему представляет решение интегрального уравнения (3.58).

3.4.5. Обобщение на комплексный случайный процесс. Распространим результаты предыдущего раздела на комплексный нормальный случайный процесс $\zeta(t)$ с нулевым средним и известной корреляционной функцией $B_\zeta(t, y)$ [см. (3.34)]. Найдем логарифм отношения правдоподобия для дискретной выборки (размером $2N$) некоррелированных координат:

$$z_k = x_k + i y_k, \quad k = 1, \ldots, N, \qquad (3.70)$$

полученных из реализации $z(t) = x(t) + iy(t)$, наблюдаемой на интервале $(-T, T)$, процесса $\zeta(t) = \xi(t) + i\eta(t)$ или процесса $\zeta(t) + s(t) = \xi(t) + a(t) + i[\eta(t) + b(t)]$, где $s(t) = a(t) + ib(t)$ — детерминированная комплексная функция. Из (3.35′) следует:

$$m_1\{x_k x_m\} = m_1\{y_k y_m\} = 0, \quad k \neq m,$$
$$m_1\{x_k y_m\} = 0 \text{ для любых } k \text{ и } m,$$
$$m_1\{x_k^2\} = m_1\{y_k^2\} = \frac{1}{2}.$$

Введем, кроме того, координаты детерминированного процесса: $s_k = a_k + ib_k$ [см. (3.30)].

Тогда в силу нормальности случайных величин x_k, y_m получим аналогично (3.62)

$$\ln l(x_1, y_1, \ldots, x_N, y_N) = \sum_{k=1}^{N} [x_k^2 + y_k^2 - (x_k - a_k)^2 - (y_k - b_k)^2]$$

или в комплексной форме

$$\ln l(z_1, \ldots, z_N) = \sum_{k=1}^{N} [|z_k|^2 - |z_k - s_k|^2] =$$

$$= 2 \sum_{k=1}^{N} \left[\operatorname{Re}(s_k \bar{z}_k) - \frac{1}{2} |s_k|^2 \right], \qquad (3.71)$$

где

$$z_k = \sqrt{\lambda_k} \int_{-T}^{T} z(t) \overline{\varphi_k(t)}\, dt; \qquad (3.72)$$

$$s_k = \sqrt{\lambda_k} \int_{-T}^{T} s(t) \overline{\varphi_k(t)}\, dt, \qquad (3.73)$$

а λ_k и $\varphi_k(t)$ — собственные числа и собственные функции интегрального уравнения (3.23) с ядром $B_\zeta(t, u)$.

Обозначим

$$V_N(t) = \sum_{k=1}^{N} \sqrt{\lambda_k} s_k \varphi_k(t). \qquad (3.74)$$

Тогда (3.71) можно переписать следующим образом:

$$\ln l(z_1, \ldots, z_N) = 2 \operatorname{Re} \int_{-T}^{T} V_N(t) \overline{\left[z(t) - \frac{s(t)}{2} \right]} dt. \qquad (3.75)$$

Если

$$\sum_{k=1}^{\infty} \lambda_k |s_k|^2 < \infty, \qquad (3.76)$$

то существует предел

$$\lim_{N \to \infty} V_N(t) = V(t),$$

который определяется из неоднородного линейного интегрального уравнения [ср. с (3.58)]

$$\int_{-T}^{T} B_\zeta(t, u) V(u)\, du = s(t), \quad |t| \leqslant T. \tag{3.77}$$

Интегральное уравнение (3.77) в комплексной форме эквивалентно системе *двух* интегральных уравнений относительно действительной $V_R(t)$ и мнимой $V_I(t)$ частей $V(t)$:

$$\int_{-T}^{T} [B_\xi(t, u) + B_\eta(t, u)] V_R(u)\, du + \int_{-T}^{T} [B_{\xi\eta}(t, u) - B_{\eta\xi}(t, u)] \times \\ \times V_I(u)\, du = \operatorname{Re} s(t), \tag{3.77'}$$

$$\int_{-T}^{T} [B_\xi(t, u) + B_\eta(t, u)] V_I(u)\, du + \\ + \int_{-T}^{T} [B_{\eta\xi}(t, u) - B_{\xi\eta}(t, u)] V_R(u)\, du = \operatorname{Im} s(t). \tag{3.77''}$$

При выполнении условия (3.76) логарифм отношения правдоподобия (3.75) сходится по вероятности к функционалу

$$\ln l\,[z(t)] = 2 \operatorname{Re} \int_{-T}^{T} \overline{V(t) \left[z(t) - \frac{1}{2} s(t) \right]}\, dt. \tag{3.78}$$

Когда ряд (3.76) расходящийся, логарифм отношения правдоподобия (3.75) стремится к $+\infty$, если реализация $z(t)$ принадлежит процессу $\zeta(t) + s(t)$, и к $-\infty$, если она принадлежит $\zeta(t)$.

3.4.6. Нормальный белый шум. Найдем в качестве примера явное выражение функционала правдоподобия, когда $\xi(t)$ представляет нормальный *белый шум* *) с интенсивностью N_0 на единицу полосы (см. § 3.4.4). Так как корреляционная функция белого шума равна

$$B(t, u) = N_0 \delta(t - u),$$

то из (3.58) сразу же находим, что

$$N_0 V(t) = s_1(t) - s_0(t), \tag{3.79}$$

---

*) См. стр. 206—207 в первой книге.

а из (3.61) находим искомое выражение для рассматриваемого случая

$$\ln l\,[x(t)] = \frac{1}{N_0}\int\limits_{-T}^{T}[s_1(t) - s_0(t)]\,x(t)\,dt -$$

$$-\frac{1}{2N_0}\int\limits_{-T}^{T}[s_1^2(t) - s_0^2(t)]\,dt. \qquad (3.80)$$

Для корреляционной функции белого шума все собственные числа интегрального уравнения (3.23) одинаковы и равны $\lambda = \frac{1}{N_0}$ (ср. с (3.24″)). Тогда условие (3.68) может быть переписано в виде

$$\sum_{k=1}^{\infty}(a_k - b_k)^2 < \infty.$$

Но сумма квадратов коэффициентов разложения $a_k - b_k$ на интервале $(-T,\ T)$ детерминированного процесса $s_0(t) - s_1(t)$ по ортогональным функциям пропорциональна (при $\lambda = \text{const}$) величине $\frac{1}{2T}\int\limits_{-T}^{T}[s_0(t) - s_1(t)]^2\,dt$.

Таким образом, при ограниченности средней мощности детерминированного процесса и при условии, что процесс $\xi(t)$ представляет белый нормальный шум, в соответствии с (3.68) *всегда будет иметь место регулярный случай.*

3.5. ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ О НОРМАЛЬНОМ СЛУЧАЙНОМ ПРОЦЕССЕ

3.5.1. Предварительное замечание. Переходим к изучению правил выбора решений о параметрах случайного процесса и их оценок, имея в виду распространение методов, изложенных в предыдущих главах, со случайных величин на случайные процессы. Когда результаты наблюдения над случайным процессом представлены выборкой (дискретной) конечного размера, все результаты, используемые для получения статистических выводов относительно случайных величин, могут быть непосредственно использованы для аналогичных целей при статистическом анализе случайных процессов. Поэтому задача последующего изло-

жения в этой главе состоит в том, чтобы показать, каким образом конструируются статистики, через которые выражаются правила выбора решений или оценки, если результатом наблюдения служит не дискретная выборка, а одна непрерывная реализация (или несколько таких реализаций) случайного процесса.

3.5.2. Проверка гипотез о среднем значении нормального случайного процесса. Начнем с простейшего случая проверки простой гипотезы H_0 о том, что среднее значение нормального случайного процесса $\xi(t)$ с известной корреляционной функцией *) $B(t, u)$ равно $s_0(t)$, против простой альтернативы H_1, что среднее равно $s_1(t)$. Будем предполагать сначала, что корреляционная функция процесса положительно определенная и, следовательно, соответствующая ей совокупность собственных функций полная (см. § 3.3.1).

Как было показано в первой главе, различные критерии качества приводят к единообразной процедуре принятия решения: по наблюденной выборке (дискретной) фиксированного размера вычисляется отношение правдоподобия и принимается или отвергается гипотеза H_0 в зависимости от того, превосходит или нет отношение правдоподобия некоторый порог, величина которого зависит от установленного заранее критерия качества и не зависит от размера выборки.

Следовательно, для проверки гипотезы о среднем значении нормального случайного процесса можно воспользоваться правилом, указанным в § 1.3, заменив отношение правдоподобия функционалом отношения правдоподобия (3.61), логарифм которого в регулярном случае ограничен по абсолютной величине. Таким образом, правило выбора решения формулируется следующим образом: если для наблюдаемой на интервале $(-T, T)$ реализации $x(t)$

$$\int_{-T}^{T} V(t)\, x(t)\, dt \geqslant \ln c + \frac{1}{2} \int_{-T}^{T} V(t)\,[s_0(t) + s_1(t)]\, dt = k_T, \quad (3.81)$$

то принимается решение γ_1 (верна гипотеза H_1), а если выполняется неравенство, противоположное (3.81), то принимается решение γ_0 (отвергается гипотеза H_1). В формуле (3.81) величина c в зависимости от выбранного крите-

*) Здесь предполагается, что $B(t, u) = m_1\{[\xi(t) - m_1\{\xi(t)\}][\xi(u) - m_1\{\xi(u)\}]\}$.

рия определяется согласно табл. 1 на стр. 40, а функция $V(t)$ является решением неоднородного линейного интегрального уравнения (3.58).

Нетрудно найти выражения условных вероятностей ошибок. Действительно, из (1.11), (1.12) и (3.81) следует:

$$\alpha = P\{\gamma_1 | H_0\} = P\left\{\int_{-T}^{T} V(t)\, x(t)\, dt \geqslant k_T \,|\, s_0(t)\right\}, \qquad (3.82)$$

$$\beta = P\{\gamma_0 | H_1\} = P\left\{\int_{-T}^{T} V(t)\, x(t)\, dt < k_T \,|\, s_1(t)\right\}. \qquad (3.83)$$

Так как $x(t)$ — реализация нормального случайного процесса, то интеграл в левой части (3.81) представляет при заданном T нормальную случайную величину, причем средние значения, вторые начальные моменты и дисперсии этой случайной величины равны:

$$m_{10} = m_1\left\{\int_{-T}^{T} V(t)\, x(t)\, dt \,|\, H_0\right\} = \int_{-T}^{T} V(t)\, s_0(t)\, dt, \qquad (3.84)$$

$$m_{11} = m_1\left\{\int_{-T}^{T} V(t)\, x(t)\, dt \,|\, H_1\right\} = \int_{-T}^{T} V(t)\, s_1(t)\, dt, \qquad (3.84')$$

$$m_1\left\{\left(\int_{-T}^{T} V(t)\, x(t)\, dt\right)^2 | H_0\right\} = m_1\left\{\int_{-T}^{T}\int_{-T}^{T} V(t)\, x(t) \times\right.$$
$$\left.\times V(u)\, x(u)\, dt\, du \,|\, H_0\right\} = \int_{-T}^{T}\int_{-T}^{T} V(t)\, V(u)\, [B(t, u) +$$
$$+ s_0(t)\, s_0(u)]\, dt\, du = \int_{-T}^{T} V(t)\, [s_1(t) - s_0(t)]\, dt +$$
$$+ \left(\int_{-T}^{T} V(t)\, s_0(t)\, dt\right)^2, \qquad (3.85)$$

$$m_1\left\{\left(\int_{-T}^{T} V(t)\, x(t)\, dt\right)^2 | H_1\right\} = \int_{-T}^{T} V(t)\, [s_1(t) - s_0(t)]\, dt +$$
$$+ \left(\int_{-T}^{T} V(t)\, s_1(t)\, dt\right)^2, \qquad (3.85')$$

т. е.

$$d_T^2 = M_2\left\{\int_{-T}^{T} V(t)\,x(t)\,dt \mid H_0\right\} =$$

$$= M_2\left\{\int_{-T}^{T} V(t)\,x(t)\,dt \mid H_1\right\} =$$

$$= \int_{-T}^{T} V(t)\,[s_1(t) - s_0(t)]\,dt = m_{11} - m_{10}. \qquad (3.86)$$

Отсюда следует, что*) [см. все формулы от (3.81) до (3.86)]

$$\alpha = \frac{1}{d_T\sqrt{2\pi}} \int_{\ln c + \frac{m_{10}+m_{11}}{2}}^{\infty} \exp\left[-\frac{(y-m_{10})^2}{2d_T^2}\right] dy =$$

$$= \int_{\frac{\ln c}{d_T}+\frac{d_T}{2}}^{\infty} \exp[-t^2/2]\,dt = 1 - F\left(\frac{\ln c}{d_T} + \frac{d_T}{2}\right), \qquad (3.87)$$

$$\beta = \frac{1}{d_T\sqrt{2\pi}} \int_{-\infty}^{\ln c + \frac{m_{10}+m_{11}}{2}} \exp\left[-\frac{(y-m_{11})^2}{2d_T^2}\right] dy =$$

$$= F\left(\frac{\ln c}{d_T} - \frac{d_T}{2}\right). \qquad (3.88)$$

Если имеется не одна, а n наблюдаемых независимых реализаций $x_i(t)$, $i = 1, \ldots, n$, то в соответствии с (3.69) область принятия гипотезы H_1 определяется неравенством

$$\int_{-T}^{T} V(t)\left[\sum_{i=1}^{n} x_i(t)\right] dt \geqslant \ln c + \frac{n}{2}\int_{-T}^{T} V(t) \times$$

$$\times [s_0(t) + s_1(t)]\,dt, \qquad (3.89)$$

---

*) Заметим также, что величины m_{10}, m_{11}, d_T, так же как и $\int_{-T}^{T} V(t)\,x(t)\,dt$, *безразмерные*, как это следует из интегрального уравнения (3.58).

а для определения условных вероятностей ошибок можно использовать те же формулы (3.87) и (3.88), полагая лишь

$$d_T^2 = n \int_{-T}^{T} V(t)\,[s_1(t) - s_0(t)]\,dt. \qquad (3.89')$$

Из указанных формул следует, что при $d_T \to \infty$ вероятности ошибок стремятся к нулю (состоятельность правила выбора решения).

Различение *близких* гипотез при заданном T требует для получения приемлемых значений вероятностей ошибок большого числа независимых реализаций. Пусть, например, $s_1(t) = a_1 s(t)$ и $s_0(t) = a_0 s(t)$. Тогда, обозначив

$$\frac{1}{a_T} = \int_{-T}^{T} V(t)\,s(t)\,dt,$$

из (3.89′) получим

$$d_T^2 = \frac{a_1 - a_0}{a_T}\,n, \qquad (3.89'')$$

откуда следует, что для сохранения *состоятельности* правила решения необходимо, чтобы при $n \to \infty$ величина $\frac{a_1 - a_0}{a_T}$ убывала медленнее чем $\frac{1}{n}$.

В том случае, когда $\xi(t)$ представляет нормальный белый шум со спектральной плотностью N_0, функция $V(t)$ определяется из (3.79). Подставляя (3.79) в (3.81), получаем следующее правило выбора решения: принимается решение γ_1, если

$$\int_{-T}^{T} [s_1(t) - s_0(t)]\,x(t)\,dt \geqslant N_0 \ln c + \frac{1}{2}\int_{-T}^{T} [s_1^2(t) - s_0^2(t)]\,dt. \qquad (3.90)$$

В частности, при равенстве энергий процессов $s_0(t)$ и $s_1(t)$ на интервале наблюдений и при $c = 1$ (критерий максимального правдоподобия) неравенство (3.90) переходит в

$$\int_{-T}^{T} s_1(t)\,x(t)\,dt \geqslant \int_{-T}^{T} s_0(t)\,x(t)\,dt. \qquad (3.90')$$

Дисперсия величины $\int\limits_{-T}^{T} V(t)\,x(t)\,dt$, когда $x(t)$ представляет реализацию суммы нормального белого шума и детерминированного процесса при любой из гипотез (H_0 и H_1), равна в соответствии с (3.86)

$$d_T^2 = \frac{1}{N_0} \int\limits_{-T}^{T} [s_1(t) - s_0(t)]^2\,dt. \qquad (3.91)$$

Покажем теперь, что в тех случаях, когда совокупность собственных функций *неполная*, можно указать правило выбора решения с нулевыми вероятностями ошибок. Как уже отмечалось, подобная ситуация (сингулярный случай) возникает, если отказаться от предположения о положительной определенности корреляционной функции $B(t, u)$ случайного процесса и допустить возможность равенства нулю квадратичной формы (3.21) или интегральной формы (3.21').

В соответствии с определением, если система ортогональных функций $\{\varphi_k(t)\}$ неполная, то существует по крайней мере еще одна функция $\psi(t)$, которая ортогональна ко всем $\varphi_k(t)$. Иначе говоря, для всех k

$$\int\limits_{-T}^{T} \varphi_k(t)\,\psi(t)\,dt = 0. \qquad (3.92)$$

Воспользуемся ортогональным разложением случайного процесса (3.30) и представим его реализацию в виде

$$x(t) = s_0(t) + \sum_{k=1}^{\infty} x_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}, \quad |t| < T,$$

когда справедлива гипотеза H_0, и в виде

$$x(t) = s_1(t) + \sum_{k=1}^{\infty} x_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}, \quad |t| \leqslant T,$$

когда справедлива гипотеза H_1.

Тогда из (3.92) следует, что

$$\int\limits_{-T}^{T} x(t)\,\psi(t)\,dt = \begin{cases} c_0, & \text{если верна гипотеза } H_0, \\ c_1, & \text{если верна гипотеза } H_1, \end{cases} \qquad (3.93)$$

где

$$c_0 = \int_{-T}^{T} s_0(t)\,\psi(t)\,dt; \quad c_1 = \int_{-T}^{T} s_1(t)\,\psi(t)\,dt. \qquad (3.93')$$

Если $c_0 \neq c_1$, то из (3.93) получаем правило выбора безошибочного решения (с вероятностью, равной единице): верна гипотеза H_0, когда

$$\int_{-T}^{T} x(t)\,\psi(t)\,dt = c_0,$$

и верна гипотеза H_1, когда

$$\int_{-T}^{T} x(t)\,\psi(t)\,dt = c_1.$$

3.5.3. Сложная альтернатива. Исследуем теперь тот случай, когда альтернатива сложная (ср. § 1.4.1). Пусть проверяется простая гипотеза H_0 о том, что среднее значение нормального случайного процесса $\xi(t)$, рассмотренного в предыдущем разделе (3.5.2), равно нулю, против сложной альтернативы H_1, что среднее принадлежит *совокупности* процессов $s_\vartheta(t)$ [или в частном случае $s(t;\vartheta)$], причем ϑ может быть любым действительным числом (или находиться в некоторых интервалах действительной оси).

Начнем с байесовского правила выбора решения. В этом случае должны быть заданы априорные вероятности q и $p = 1 - q$ того, что $m_1\{\xi(t)\} = 0$ и $m_1\{\xi(t)\} = s_\vartheta(t)$, и априорное распределение $w_1(\vartheta)$ параметра ϑ. Усредняя отношение правдоподобия по случайному параметру ϑ, для дискретной выборки некоррелированных координат $x_1, \ldots, x_N$ находим [см. (1.110′) и (3.62)]

$$\Lambda(x_1, \ldots, x_N) = \int_{-\infty}^{\infty} \exp\left\{\sum_{k=1}^{N}\left[x_k b_k(\vartheta) - \right.\right.$$

$$\left.\left. - \frac{1}{2} b_k^2(\vartheta)\right]\right\} w_1(\vartheta)\,d\vartheta, \qquad (3.94)$$

где

$$x_k = \sqrt{\lambda_k} \int_{-T}^{T} x(t)\,\varphi_k(t)\,dt; \qquad (3.95)$$

$$b_k(\vartheta) = \sqrt{\lambda_k} \int_{-T}^{T} s_\vartheta(t)\,\varphi_k(t)\,dt; \qquad (3.95')$$

λ_k, $\varphi_k(t)$ — собственные числа и собственные функции однородного линейного интегрального уравнения (3.23).

Аналогично предыдущему, если $\sum_{k=1}^{\infty} \lambda_k b_k^2(\vartheta)$ конечна, то существует предел

$$\lim_{N\to\infty} \sum_{k=1}^{N} \sqrt{\lambda_k}\, b_k(\vartheta)\,\varphi_k(t) = V(t;\ \vartheta), \qquad (3.96)$$

причем функция $V(t;\ \vartheta)$ определяется из неоднородного линейного интегрального уравнения

$$\int_{-T}^{T} B(t,\ u)\,V(u;\ \vartheta)\,du = s_\vartheta(t),\ |t| \leqslant T. \qquad (3.97)$$

При этих условиях усредненное отношение правдоподобия сходится по вероятности к функционалу

$$\Lambda[x(t)] = \int_{-\infty}^{\infty} \exp\left\{\int_{-T}^{T} V(t;\ \vartheta)\left[x(t) - \frac{s_\vartheta(t)}{2}\right] dt\right\} w_1(\vartheta)\,d\vartheta. \qquad (3.98)$$

Используя (3.98), приходим к следующему байесовскому правилу (минимизирующему средний риск): принимается решение γ_1 (отвергается гипотеза H_0), если для наблюдаемой реализации

$$\int_{-\infty}^{\infty} \exp\left\{\int_{-T}^{T} V(t;\ \vartheta)\left[x(t) - \frac{s_\vartheta(t)}{2}\right] dt\right\} w_1(\vartheta) \geqslant c = \frac{q}{p}\,\frac{\Pi_{01} - \Pi_{00}}{\Pi_{10} - \Pi_{11}}, \qquad (3.99)$$

и решение γ_0 (принимается гипотеза H_0), если выполняется неравенство, противоположное (3.99) [см. (1.109)].

В частном случае, когда $s_\vartheta(t) = a(\vartheta)\, s(t)$ — квазидетерминированный процесс со случайной амплитудой, функционал (3.98) отношения правдоподобия может быть переписан в виде

$$\Lambda[x(t)] = \int_{-\infty}^{\infty} \exp\left[a(\vartheta) \int_{-T}^{T} x(t)\, V(t)\, dt - \right.$$

$$\left. - \frac{a^2(\vartheta)}{2} \int_{-T}^{T} s(t)\, V(t)\, dt\right] w_1(\vartheta)\, d\vartheta, \qquad (3.99')$$

где $V(t)$ — решение интегрального уравнения

$$\int_{-T}^{T} B(t, u)\, V(u)\, du = s(t), \quad |t| \leqslant T. \qquad (3.100)$$

Рассмотрим теперь критерий Неймана — Пирсона. Повторяя рассуждения, приведенные в начале этого раздела, получим при произвольных ϑ следующее выражение логарифма функционала отношения правдоподобия:

$$\ln l[x(t)] = \int_{-T}^{T} V(t;\ \vartheta)\, x(t)\, dt -$$

$$- \frac{1}{2} \int_{-T}^{T} V(t;\ \vartheta)\, s_\vartheta(t)\, dt. \qquad (3.101)$$

Критическая область для фиксированного ϑ, в которой отклоняется гипотеза H_0, определяется неравенством

$$\int_{-T}^{T} V(t;\ \vartheta)\, x(t)\, dt \geqslant K(\vartheta), \qquad (3.102)$$

где $K(\vartheta)$ находится по заданной величине вероятности ошибки первого рода. Из (3.102) следует, что для существования равномерно наиболее мощного правила (см. § 1.4.4 и § 1.4.7) должно выполняться условие

$$s_\vartheta(t) = a(\vartheta)\, s(t), \qquad (3.103)$$

причем амплитуда $a(\vartheta)$ сохраняет постоянный знак. При условии $a(\vartheta) > 0$ правило выбора решения формулируется так: отвергается гипотеза H_0, если для наблюденной реали-

зации $x(t)$

$$\int_{-T}^{T} V(t)\,x(t)\,dt \geqslant K, \qquad (3.104)$$

где $V(t)$ определяется из интегрального уравнения (3.100).

Так как при справедливости гипотезы H_0

$$m_1\left\{\int_{-T}^{T} V(t)\,x(t)\,dt\right\} = 0,$$

$$M_2\left\{\int_{-T}^{T} V(t)\,x(t)\,dt\right\} = \int_{-T}^{T} V(t)\,s(t)\,dt = d_T^2,$$

то в силу нормального распределения интеграла в (3.104)

$$P\left\{\int_{-T}^{T} V(t)\,x(t)\,dt \geqslant K \mid H_0\right\} = 1 - F\left(\frac{K}{d_T}\right) \qquad (3.105)$$

и, следовательно, при данной вероятности α ошибки первого рода [ср. с (1.76)]

$$K = d_T x_\alpha, \qquad (3.105')$$

где x_α — процентная точка нормального распределения.

Если $a(\vartheta) < 0$, то неравенство (3.104) заменяется следующим:

$$\int_{-T}^{T} V(t)\,x(t)\,dt < -K. \qquad (3.106)$$

Итак, правило (3.104) является равномерно наиболее мощным при сложной альтернативе, для которой $a(\vartheta) > 0$, а (3.106) — равномерно наиболее мощным при альтернативе, для которой $a(\vartheta) < 0$. Если же рассматривать в качестве сложной альтернативы все действительные значения амплитуды $a(\vartheta)$, то равномерно наиболее мощного правила не существует.

Функция мощности для правила (3.104) имеет вид

$$1 - \beta(\vartheta) = P\left\{\int_{-T}^{T} V(t)\,x(t)\,dt \geqslant K \mid H_1\right\} =$$

$$= 1 - F\left(\frac{K}{d_T} - a(\vartheta)\,d_T\right). \qquad (3.107)$$

При $a(\vartheta) > 0$, $1 - \beta(\vartheta) > \alpha$, функция мощности монотонно возрастает при увеличении $a(\vartheta)$. Но если $a(\vartheta) < 0$, то $1 - \beta(\vartheta) < \alpha$ и функция мощности монотонно стремится к нулю при неограниченном увеличении $|a(\vartheta)|$. Следовательно, если область возможных значений $a(\vartheta)$ — все действительные числа, то правило (3.104) смещенное.

Можно показать [2], что аналогично дискретному случаю, рассмотренному в § 1.4.7, равномерно наиболее мощное несмещенное правило определяется по наблюдаемой реализации следующим соотношением:

$$\left| \int_{-T}^{T} x(t) V(t) \, dt \right| \geqslant K. \qquad (3.108)$$

3.5.4. Процессы с дробно-рациональными энергетическими спектрами. Из изложенного ясно, что основная сложность полученных правил выбора решения для проверки гипотезы о среднем значении нормального случайного процесса связана с тем, что зависимость правила от известной корреляционной функции $B(t, u)$ процесса выражена не в явной форме, а через решение $V(t)$ неоднородного линейного интегрального уравнения (3.58). Решение этого уравнения в общем виде неизвестно, в некоторых случаях оно может быть получено ценой весьма громоздких построений, и всегда остается путь для использования численных методов. Исключением является случай, когда нормальный процесс представляет сумму детерминированной функции и стационарного нормального случайного процесса $\xi(t)$ с нулевым средним и дробно-рациональным энергетическим спектром $F_\xi(\omega)$ вида

$$F_\xi(\omega) = \frac{Q(\omega)}{P(\omega)},$$

где Q и P — многочлены степени m и n соответственно ($m < n$). При этом процесс $\xi(t)$ является *компонентой многомерного* нормального марковского процесса (см. § 9.5.1 в первой книге). В указанном случае интегральное уравнение (3.58) сводится к обыкновенному дифференциальному уравнению с постоянными коэффициентами и его решение может быть записано в явном виде (см., например, [3, приложение II]).

Рассмотрим задачу проверки простой гипотезы H_0 о том, что наблюдаемая реализация принадлежит процес-

су $\xi(t)+s_0(t)$, против простой альтернативы H_1, что она принадлежит $\xi(t)+s_1(t)$, при условии, что $\xi(t)$ — нормальный стационарный марковский процесс. Тогда, как отмечалось в § 4.5.4 первой книги, энергетический спектр процесса равен

$$F_\xi(\omega)=\frac{\sigma^2}{\mu^2+\omega^2},$$

а его корреляционная функция

$$B(\tau)=\sigma^2 e^{-\mu|\tau|}, \quad \mu>0. \tag{3.190}$$

Решение интегрального уравнения (3.58) с ядром типа (3.109) имеет следующий вид:

$$V(t)=\frac{1}{2\mu\sigma^2}\{[\mu s_\Delta(-T)-s'_\Delta(-T)]\,\delta(t+T)+$$

$$+[\mu s_\Delta(T)-s'_\Delta(T)]\,\delta(t-T)+\mu^2 s_\Delta(t)-s''_\Delta(t)\}, \qquad |t|\leqslant T, \tag{3.110}$$

где

$$s_\Delta(t)=s_1(t)-s_0(t). \tag{3.110$'$}$$

Подставляя (3.110′) в (3.81) и вводя обозначение для суммы детерминированных функций

$$s_\Sigma(t)=s_1(t)+s_0(t), \tag{3.110$''$}$$

запишем правило выбора решения в рассматриваемом случае следующим образом: принимается решение γ_1 (справедлива гипотеза H_1), если для наблюдаемой на интервале $(-T, T)$ реализации $x(t)$ процесса

$$\left[s_\Delta(-T)-\frac{s'_\Delta(-T)}{\mu}\right]x(-T)+\left[s_\Delta(T)-\frac{s'_\Delta(T)}{\mu}\right]x(T)+$$

$$+\mu\int_{-T}^{T}x(t)\left[s_\Delta(t)-\frac{s''_\Delta(t)}{\mu^2}\right]dt>2\sigma^2\ln c+\frac{1}{2}\left\{\left[s_\Delta(-T)-\right.\right.$$

$$\left.-\frac{s'_\Delta(-T)}{\mu}\right]s_\Sigma(-T)+\left[s_\Delta(T)-\frac{s'_\Delta(T)}{\mu}\right]s_\Sigma(T)+\mu\int_{-T}^{T}s_\Sigma(t)\times$$

$$\left.\times\left[s_\Delta(t)-\frac{s''_\Delta(t)}{\mu^2}\right]dt\right\}. \tag{3.111}$$

Условные вероятности ошибок первого и второго рода определяются величиной [см. (3.87) и (3.88)]

$$d_T^2 = \frac{1}{2\sigma^2}\left\{\left[s_\Delta(-T) - \frac{s'_\Delta(-T)}{\mu}\right] s_\Delta(-T) + \right.$$

$$+\left[s_\Delta(T) - \frac{s'_\Delta(T)}{\mu}\right] s_\Delta(T) + \mu\int_{-T}^{T}\left[s_\Delta(t) - \frac{s''_\Delta(t)}{\mu^2}\right] s_\Delta(t)\,dt\right\} =$$

$$= \frac{1}{2\sigma^2}\left\{s_\Delta^2(-T) + s_\Delta^2(T) + \mu\int_{-T}^{T}\left(s_\Delta^2(t) + \frac{1}{\mu^2}[s'_\Delta(t)]^2\right)dt\right\}. \tag{3.112}$$

Если при $\mu T \gg 1$ величины $\frac{s_\Delta(|T|)}{\sigma} < 1$ и $\frac{s'_\Delta(|T|)}{\mu\sigma} < 1$, то из (3.111) и (3.112) получаем следующие асимптотические выражения, справедливые при большом времени наблюдения:

$$\mu\int_{-T}^{T} x(t)\left[s_\Delta(t) - \frac{s''_\Delta(t)}{\mu^2}\right]dt \geqslant 2\sigma^2 \ln c +$$

$$+\frac{\mu}{2}\int_{-T}^{T} s_\Sigma(t)\left[s_\Delta(t) - \frac{s''_\Delta(t)}{\mu^2}\right]dt, \tag{3.113}$$

$$d_T^2 \sim \frac{\mu}{2\sigma^2}\int_{-T}^{T}\left\{s_\Delta^2(t) + \frac{1}{\mu^2}[s'_\Delta(t)]^2\right\}dt. \tag{3.114}$$

В том случае, когда для любых t функции $s_0(t)$ и $s_1(t)$ постоянны, т. е.

$$s_0(t) \equiv a_0, \quad s_1(t) \equiv a_1, \tag{3.115}$$

рассмотренная задача сводится к проверке гипотезы о величине среднего значения стационарного нормального случайного процесса с корреляционной функцией вида (3.109). Тогда производные от функций $s_0(t)$ и $s_1(t)$ тождественно равны нулю, и из (3.111) следует правило: среднее равно a_1, если

$$x(-T) + x(T) + \mu\int_{-T}^{T} x(t)\,dt \geqslant \frac{2\sigma^2}{a_1 - a_0}\ln c +$$

$$+(a_1 + a_0)(1 + \mu T), \tag{3.116}$$

При этом

$$d_T^2 = \frac{(a_1 - a_0)^2}{\sigma^2}(1 + \mu T). \qquad (3.117)$$

Если $\mu T \gg 1$, то имеем асимптотические соотношения

$$\frac{x(-T) + x(T)}{2\mu T} + \frac{1}{2T}\int\limits_{-T}^{T} \quad (t)\,dt \gtrless \frac{a_1 + a_0}{2}, \qquad (3.118)$$

а величина d_T^2 возрастает пропорционально μT:

$$d_T^2 \sim \left(\frac{a_1 - a_0}{\sigma}\right)^2 \mu T. \qquad (3.118')$$

3.5.5. Проверка гипотез о корреляционной функции. Проверяется гипотеза H_0 о том, что корреляционная функция нормального случайного процесса $\xi(t)$ с нулевым средним *) равна $B_0(t, y)$, против альтернативы H_1, что его корреляционная функция равна $B_1(t, y)$. Если выбрать в качестве координат процесса $\xi(t)$ случайные величины [см. (3.25)]

$$\xi_k = \sqrt{\lambda_k}\int\limits_{-T}^{T} \xi(t)\,\varphi_k(t)\,dt, \qquad (3.119)$$

где λ_k и $\varphi_k(t)$ — собственные числа и собственные функции интегрального уравнения

$$\varphi(t) = \lambda\int\limits_{-T}^{T} B_0(t, u)\,\varphi(u)\,du, \quad |t| \leqslant T,$$

то эти координаты не коррелированы, если справедлива гипотеза H_0, но будут *коррелированы*, если верна гипотеза H_1.

Можно, однако, выбрать координаты процесса таким образом, чтобы они были не коррелированными (а, следовательно, независимы в силу нормального распределения) и при гипотезе H_0, и при гипотезе H_1, с той лишь разницей, что при одной гипотезе дисперсии всех координат единичны, а при другой — различны для разных координат. Здесь имеет место аналогия с известным результатом высшей

---

*) Легко обобщить результаты данного раздела, предположив, что среднее значение процесса отлично от нуля [12]. Дискретным аналогом рассматриваемой задачи является задача 1.6.

алгебры, согласно которому одним линейным преобразованием можно одну квадратичную форму привести к нормальному виду (т. е. к сумме квадратов переменных), а другую — к каноническому *).

Пусть координаты процесса $\xi(t)$ определяются согласно (3.119), причем λ_k и $\varphi_k(t)$ — собственные числа и собственные функции (ненормированные) интегрального уравнения

$$\int_{-T}^{T} [B_1(t, u) - \lambda B_0(t, u)]\, \varphi(u)\, du = 0, \quad |t| \leqslant T, \quad (3.120)$$

а нормировка собственных функций производится относительно $B_0(t, y)$, т. е.

$$\lambda_k \int_{-T}^{T} \int_{-T}^{T} B_0(t, u)\, \varphi_k(t)\, \varphi_m(u)\, dt\, du = \begin{cases} 1, & k = m, \\ 0, & k \neq m. \end{cases} \quad (3.120')$$

Тогда

$$m_1\{\xi_k \xi_m \mid H_0\} = \sqrt{\lambda_k \lambda_m} \int_{-T}^{T} \int_{-T}^{T} B_0(t, u)\, \varphi_k(t)\, \varphi_m(u)\, dt\, du = \begin{cases} 1, & k = m, \\ 0, & k \neq m, \end{cases} \quad (3.121)$$

$$m_1\{\xi_k \xi_m \mid H_1\} = \sqrt{\lambda_k \lambda_m} \int_{-T}^{T} \int_{-T}^{T} B_1(t, u) \times \varphi_k(t)\, \varphi_m(u)\, dt\, du = \lambda_m \sqrt{\lambda_k \lambda_m} \int_{-T}^{T} \int_{-T}^{T} B_0(t, u) \times \varphi_k(t)\, \varphi_m(u)\, dt\, du = \begin{cases} \lambda_k, & k = m, \\ 0, & k \neq m. \end{cases} \quad (3.121')$$

Используя (3.121) и (3.121′), нетрудно записать логарифм отношения правдоподобия для N наблюдаемых независимых координат $x_1, \ldots, x_N$ реализации нормального

---

*) См., например: Мишина А. П. и Проскуряков И. В. Высшая алгебра. Физматгиз, 1962, стр. 162—164.

стационарного случайного процесса:

$$\ln l(x_1, \ldots, x_N) = -\frac{1}{2}\sum_{k=1}^{N}\ln\lambda_k - \frac{1}{2}\sum_{k=1}^{N}\left(\frac{x_k^2}{\lambda_k} - x_k^2\right) = \frac{1}{2}\sum_{k=1}^{N}\frac{\lambda_k - 1}{\lambda_k}x_k^2 - \frac{1}{2}\sum_{k=1}^{N}\ln\lambda_k, \qquad (3.122)$$

где в соответствии с (3.119)

$$x_k = \sqrt{\lambda_k}\int_{-T}^{T} x(t)\,\varphi_k(t)\,dt, \qquad (3.122')$$

причем λ_k и $\varphi_k(t)$ — собственные числа и собственные функции интегрального уравнения (3.120), а $x(t)$ — реализация нормального случайного процесса, наблюдаемая на интервале $(-T, T)$.

Подставляя (3.122′) в (3.122) и переходя к пределу при $N \to \infty$, получим следующее выражение логарифма функционала отношения правдоподобия:

$$\ln l[x(t)] = \frac{1}{2}\sum_{k=1}^{\infty}(\lambda_k - 1)\times \left[\int_{-T}^{T} x(t)\,\varphi_k(t)\,dt\right]^2 - \frac{1}{2}\sum_{k=1}^{\infty}\ln\lambda_k. \qquad (3.123)$$

Теперь может быть сформулировано правило выбора решения. Если для наблюдаемой на интервале $(-T, T)$ реализации $x(t)$

$$\sum_{k=1}^{\infty}(\lambda_k - 1)\left[\int_{-T}^{T} x(t)\,\varphi_k(t)\,dt\right]^2 \geqslant 2\ln c + \sum_{k=1}^{\infty}\ln\lambda_k, \qquad (3.124)$$

то принимается решение γ_1 (верна гипотеза H_1), а если выполняется неравенство, противоположное (3.124), то принимается решение γ_0 (верна гипотеза H_0). Порог c, как всегда, выбирается в соответствии с принятым критерием качества (см. табл. 1 на стр. 40).

Для того чтобы найти вероятности ошибок, необходимо предварительно определить функцию распределения логарифма функционала отношения правдоподобия (3.123).

Хотя $\ln l\,[x(t)]$ представляет бесконечную сумму независимых случайных величин $\frac{\lambda_k-1}{2\lambda_k}x_k^2-\frac{1}{2}\ln\lambda_k$, распределение этой суммы отличается от *нормального*, так как условие применимости предельной теоремы Ляпунова не выполняется [см. (3.144) в первой книге]. Для определения указанного распределения воспользуемся методом характеристических функций. Так как x_k — нормальная случайная величина, то одномерные характеристические функции k-го слагаемого суммы при гипотезах H_0 и H_1 равны соответственно [см. (9.16) в первой книге]

$$\Theta_{1k}(v\,|\,H_0)=\frac{1}{\sqrt{1-iv\left(1-\frac{1}{\lambda_k}\right)}}\,e^{-\frac{iv}{2}\ln\lambda_k}, \qquad (3.125)$$

$$\Theta_{1k}(v\,|\,H_1)=\frac{1}{\sqrt{1-iv(\lambda_k-1)}}\,e^{-\frac{iv}{2}\ln\lambda_k}. \qquad (3.125')$$

Учитывая независимость случайных величин x_k, находим одномерные характеристические функции логарифма функционала отношения правдоподобия как произведения характеристических функций (3.125) и (3.125′):

$$\Theta_1(v\,|\,H_0)=\left[\prod_{k=1}^{\infty}\left(1-\frac{iv}{1+v_k}\right)\right]^{-\frac{1}{2}}\times$$
$$\times\exp\left\{-\frac{iv}{2}\ln\prod_{k=1}^{\infty}\left(1+\frac{1}{v_k}\right)\right\}, \qquad (3.126)$$

$$\Theta_1(v\,|\,H_1)=\left[\prod_{k=1}^{\infty}\left(1-\frac{iv}{v_k}\right)\right]^{-\frac{1}{2}}\times$$
$$\times\exp\left\{-\frac{iv}{2}\ln\prod_{k=1}^{\infty}\left(1+\frac{1}{v_k}\right)\right\}, \qquad (3.126')$$

где обозначено

$$v_k=\frac{1}{\lambda_k-1}. \qquad (3.126'')$$

Теперь обратным преобразованием Фурье можно найти плотности вероятности и интегральные функции распределения логарифма функционала отношения правдоподобия. Однако эти вычисления представляют в общем случае довольно сложную математическую задачу.

Остановимся еще на том случае, когда коэффициент корреляции $R(\tau)$ стационарного нормального случайного процесса точно известен и проверяется лишь гипотеза H_0 о том, что дисперсия равна σ_0^2, против альтернативы H_1, что она равна σ_1^2. В рассматриваемом случае $B_0(\tau) = \sigma_0^2 R(\tau)$, $B_1(\tau) = \sigma_1^2 R(\tau)$. Обозначим $\left(\frac{\sigma_1}{\sigma_0}\right)^2 = \mu^2$, и пусть $\mu \neq 1$. Тогда из (3.120) находим

$$(\mu^2 - \lambda) \int_{-T}^{T} R(t-u)\,\varphi(u)\,du = 0, \qquad (3.127)$$

откуда следует, что собственные числа интегрального уравнения постоянны и равны $\lambda_k = \mu^2$. Логарифм отношения правдоподобия в соответствии с (3.122) преобразуется к виду

$$\ln l(x_1, \ldots, x_N) = \frac{1}{2}\left(1 - \frac{1}{\mu^2}\right) \sum_{k=1}^{N} x_k^2 - \frac{N}{2} \ln \mu^2,$$

и, следовательно:

$$\lim_{N\to\infty} \frac{1}{N} \ln l(x_1, \ldots, x_N) = \frac{1}{2}\left(1 - \frac{1}{\mu^2}\right) \times$$

$$\times \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} x_k^2 - \ln \mu =$$

$$= \begin{cases} \frac{1}{2}\left(1 - \frac{1}{\mu^2}\right) - \ln \mu < 0, & \text{если верна гипотеза } H_0, \\ \frac{1}{2}(\mu^2 - 1) - \ln \mu > 0, & \text{если верна гипотеза } H_1, \end{cases} \qquad (3.128)$$

так как последовательность случайных величин $z_N = \frac{1}{N} \sum_{k=1}^{N} x_k^2$ сходится по вероятности к единице *), если

верна гипотеза H_0, и к μ^2, если верна гипотеза H_1. Из (3.128) следует, что

$$\lim_{N\to\infty} \ln l(x_1, \ldots, x_N) = -\infty, \tag{3.129}$$

если верна гипотеза H_0, и

$$\lim_{N\to\infty} \ln l(x_1, \ldots, x_N) = \infty, \tag{3.129'}$$

если верна гипотеза H_1.

Таким образом, рассматриваемый случай сингулярный, и поэтому возможно при любом (произвольно малом) времени наблюдения выбрать правило проверки гипотезы с вероятностью единица. Такое правило непосредственно следует из (3.128), если вместо координат x_k подставить их выражения (3.122') через наблюдаемую реализацию. Если для наблюдаемой на интервале $(-T, T)$ реализации $x(t)$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \lambda_k \left(\int_{-T}^{T} x(t)\, \varphi_k(t)\, dt \right)^2 = 1, \tag{3.130}$$

то принимается гипотеза H_0 (дисперсия процесса равна σ_0^2, а если предел в (3.130) равен μ^2, то принимается гипотеза H_1 (дисперсия процесса равна σ_1^2).

3.5.6. Случай, когда проверяемая гипотеза — белый шум. Рассмотрим частный случай общей постановки задачи, изложенной в предыдущем разделе. Пусть проверяется гипотеза о том, что $\xi(t)$ представляет нормальный белый шум со спектральной интенсивностью N_0, против альтернативы, что корреляционная функция нормального случайного процесса равна $B(t, y) = N_0\delta(t-y)$. Покажем, что в этом

---

*) Заметим, что для суммы z_N выполняются условия применимости теоремы Ляпунова, так что при $N \to \infty$

$$W_1(z_N) \sim \frac{1}{\sqrt{2\pi M_2\{z_N\}}} \exp\left\{ -\frac{1}{2M_2\{z_N\}} [z_N - m_1\{z_N\}]^2 \right\},$$

причем $m_1\{z_N \mid H_0\} = 1$, $m_1\{z_N \mid H_1\} = \mu^2$, $M_2\{z_N \mid H_0\} = \frac{2}{N}$, $M_2\{z_N \mid H_1\} = \frac{2\mu^2}{N}$. В пределе $\lim\limits_{N\to\infty} W_1(z_N) = \delta(z - m_1\{z\})$.

случае в (3.124) суммирование может быть заменено интегрированием реализаций с некоторой весовой функцией, зависящей от $B(t, y)$.

В рассматриваемом случае интегральное уравнение (3.120) преобразуется к виду

$$N_0(\lambda - 1)\varphi(t) = \int_{-T}^{T} B(t, y)\varphi(y)\,dy, \quad |t| < T, \qquad (3.131)$$

причем для нормировки каждую из собственных функцией $\varphi_k(t)$ необходимо разделить на $\sqrt{\lambda_k N_0}$ [см. (3.120′)]. Обозначим

$$h(t, y) = \sum_{k=1}^{\infty} \frac{\lambda_k - 1}{\lambda_k}\varphi_k(t)\varphi_k(y), \qquad (3.132)$$

где $\varphi_k(t)$ — ортогональны и *нормированы*. Тогда первое слагаемое в (3.122) может быть представлено в виде

$$\sum_{k=1}^{\infty} \frac{\lambda_k - 1}{\lambda_k} x_k^2 = \int_{-T}^{T}\int_{-T}^{T} x(t)\,x(y) \sum_{k=1}^{\infty} \frac{\lambda_k - 1}{\lambda_k N_0} \times$$

$$\times \varphi_k(t)\varphi_k(y)\,dt\,dy = \frac{1}{N_0}\int_{-T}^{T}\int_{-T}^{T} h(t, y)\,x(t)\,x(y)\,dt\,dy. \qquad (3.133)$$

Далее из (3.132) находим

$$\int_{-T}^{T} [B(t, y) + N_0\delta(t - y)]\,h(v, y)\,dy = \sum_{k=1}^{\infty} \frac{\lambda_k - 1}{\lambda_k}\varphi_k(v) \times$$

$$\times \int_{-T}^{T} [B(t, y) + N_0\delta(t - y)]\,\varphi_k(y)\,dy = \sum_{k=1}^{\infty} N_0(\lambda_k - 1)\varphi_k(v) \times$$

$$\times \varphi_k(t) = B(t, v).$$

Следовательно, весовая функция $h(t, u)$ удовлетворяет интегральному уравнению

$$\int_{-T}^{T} B(t, y)\,h(v, y)\,dy + N_0 h(v, t) =$$

$$= B(t, v), \quad |t| \leqslant T, \ |v| \leqslant T. \qquad (3.134)$$

Интегральное уравнение (3.134) является частным случаем (при $u=1$) более общего уравнения

$$u\int_{-T}^{T} B(t,y)\,h(v,t,u)\,dy+N_0h(v,t,u)=B(t,v),$$
$$|t|\leqslant T,\ |v|\leqslant T, \qquad (3.135)$$

относительно функции $h(v,t,u)$, которую называют *резольвентой*. Для резольвенты имеет место разложение

$$h(t,y,u)=\sum_{k=1}^{\infty}\frac{\lambda_k-1}{1+(\lambda_k-1)u}\varphi_k(t)\varphi_k(y), \qquad (3.136)$$

из которого при $u=1$ следует (3.132). Из (3.136) находим

$$\int_{-T}^{T} h(t,t,u)\,dt=\sum_{k=1}^{\infty}\frac{\lambda_k-1}{1+(\lambda_k-1)u}, \qquad (3.136')$$

откуда интегрированием по переменной u от нуля до единицы получаем

$$\int_0^1\int_{-T}^{T} h(t,t,u)\,dt\,du=\int_0^1\left(\sum_{k=1}^{\infty}\frac{\lambda_k-1}{1+(\lambda_k-1)u}\right)du=$$
$$=\sum_{k=1}^{\infty}\int_0^{\lambda_k-1}\frac{dx}{1+x}=\sum_{k=1}^{\infty}\ln\lambda_k.$$

Таким образом, второе слагаемое в правой части (3.123) в рассматриваемом случае равно

$$\sum_{k=1}^{\infty}\ln\lambda_k=\int_0^1\int_{-T}^{T} h(t,t,u)\,dt\,du. \qquad (3.137)$$

Подставляя (3.133) и (3.137) в (3.124), получим следующее правило для проверки гипотезы H_0 о том, что наблюдаемая реализация принадлежит нормальному белому шуму: гипотеза H_0 отвергается, если

$$\frac{1}{N_0}\int_{-T}^{T}\int_{-T}^{T} h(t,y)\,x(t)\,x(y)\,dt\,dy\geqslant 2\ln c+$$
$$+\int_0^1\int_{-T}^{T} h(t,t,u)\,dt\,du, \qquad (3.138)$$

где $h(t, y)$ и $h(t, y, u)$ — решения интегральных уравнений (3.134) и (3.135) соответственно.

Для определения в рассматриваемом случае характеристических функций логарифма отношения правдоподобия воспользуемся формулами (3.126), (3.126′). Для этого введем сначала функцию $\ln D(z)$, равную интегралу от (3.136′) по u в пределах от нуля до z, т. е.

$$\ln D(z) = \int_0^z \int_{-T}^{T} h(t, t, u)\, dt\, du =$$

$$\sum_{k=1}^{\infty} \ln\left(1 + \frac{z}{v_k}\right) = \ln \prod_{k=1}^{\infty}\left(1 + \frac{z}{v_k}\right) \qquad (3.139)$$

или

$$D(z) = \exp\left\{\int_0^z \int_{-T}^{T} h(t, t, u)\, dt\, du\right\} = \prod_{k=1}^{\infty}\left(1 + \frac{z}{v_k}\right), \qquad (3.139')$$

где v_k определено согласно (3.126″).

Используя (3.139) и (3.139′), получаем искомые выражения характеристических функций [см. (3.126), (3.126′)]

$$\Theta_1(v \mid H_0) = \sqrt{\frac{D(1)}{D(1-iv)}} \exp\left\{-\frac{iv}{2} \ln D(1)\right\}, \qquad (3.140)$$

$$\Theta_1(v \mid H_1) = \frac{1}{\sqrt{D(-iv)}} \exp\left\{-\frac{iv}{2} \ln D(1)\right\}. \qquad (3.140')$$

Тогда условные вероятности ошибок первого и второго рода при использовании правила (3.138) равны

$$\alpha = \frac{\sqrt{D(1)}}{2\pi} \int_{\ln c}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{D(1-iv)}} \exp\left\{-iv\left[x + \ln \sqrt{D(1)}\right]\right\} dv\, dx, \qquad (3.141)$$

$$\beta = \frac{1}{2\pi} \int_{-\infty}^{\ln c} \int_{-\infty}^{\infty} \frac{1}{\sqrt{D(-iv)}} \exp\left\{-iv\left[x + \ln \sqrt{D(1)}\right]\right\} dv\, dx. \qquad (3.141')$$

3.5.7. Обобщение для процессов с дробно-рациональными энергетическими спектрами. Можно распространить результаты предыдущего раздела на тот случай, когда

проверяемая гипотеза и альтернатива относятся к нормальным случайным процессам с дробно-рациональными энергетическими спектрами (см. § 3.5.4). Для того чтобы сохранялось соотношение (3.133), весовая функция $h(t, y)$ должна удовлетворять двумерному интегральному уравнению

$$\frac{1}{N_0}\int_{-T}^{T}\int_{-T}^{T} h(s, y) B_0(t-s) B_1(y-v)\,ds\,dy = B_1(t-v)-B_0(t-v), \quad |t|\leq T,\ |v|\leq T, \tag{3.142}$$

где N_0 — постоянный (размерный) множитель. При $B_1(t-v) = B(t-v) + N_0\delta(t-v)$, $B_0(t-v) = N_0\delta(t-v)$ уравнение (3.142) переходит в (3.134) [с дополнительным ограничением стационарности процесса и рациональности спектра $F_1(\omega)$].

Вводя двумерное преобразование Фурье от функции $\frac{1}{N_0}h(s, y)$ по формуле

$$\frac{1}{N_0}h(s, y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(\omega, \lambda)\,e^{i(\lambda s-\omega y)}\,d\lambda\,d\omega,$$

можно интегральное уравнение (3.142) переписать в виде (см. [13])

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(\omega, \lambda) F_0(\omega) F_1(\lambda)\,e^{-i(\lambda t-\omega v)}\,d\omega\,d\lambda = B_1(t-v)-B_0(t-v), \quad |t|\leq T,\ |v|\leq T, \tag{3.142'}$$

где $F_0(\omega)$ и $F_1(\omega)$ — энергетические спектры двух стационарных нормальных случайных процессов, каждый из которых представляется в виде

$$F(\omega) = \left|\frac{G(i\omega)}{V(i\omega)}\right|^2, \tag{3.143}$$

причем все корни полиномов $G(i\omega)$ и $V(i\omega)$ имеют отрицательные действительные части, и степень полинома числителя не выше степени полинома знаменателя.

Обозначая в (3.142') интеграл по переменной ω

$$L(\lambda, v) = \int_{-\infty}^{\infty} H(\omega, \lambda) F_0(\omega)\,e^{i\omega v}\,d\omega, \tag{3.144}$$

приводим интегральное уравнение (3.142′) к виду

$$\int_{-\infty}^{\infty} L(\lambda, v) F_1(\omega) e^{-i\lambda t} d\lambda = B_1(t-v) - B_0(t-v),$$

$$|t| \leqslant T, \ |v| \leqslant T. \qquad (3.145)$$

Уравнения (3.144) и (3.145) представляют два однотипных неоднородных линейных интегральных уравнения, каждое из которых сводится к решению линейного дифференциального уравнения (см. [13]), причем сначала решается уравнение (3.145), полагая параметром переменную v. Затем находится решение уравнения (3.144) относительно $H(\omega, \lambda)$, полагая параметром переменную λ.

Общее правило выбора решения (3.124) приводится в этом случае к виду, аналогичному (3.138): отвергается гипотеза H_0, если

$$\frac{1}{N_0} \int_{-T}^{T} \int_{-T}^{T} h(t, y) x(t) x(y) dt\, dy \geqslant 2 \ln c + \sum_{k=1}^{\infty} \ln \lambda_k, \qquad (3.146)$$

причем первое слагаемое в правой части (3.146) определяется, как всегда, критерием качества, а определение второго слагаемого сопряжено с дополнительными достаточно громоздкими вычислениями. Как показано в [13], величины $\delta_k = \frac{\lambda_k}{1-\lambda_k}$ являются собственными значениями дифференциального уравнения (записанного в операторной форме)

$$G_1\left(\frac{d}{dt}\right) G_1\left(-\frac{d}{dt}\right) V_2\left(\frac{d}{dt}\right) V_2\left(-\frac{d}{dt}\right) y(t) =$$
$$= \left(1 + \frac{1}{\delta}\right) G_2\left(\frac{d}{dt}\right) G_2\left(-\frac{d}{dt}\right) V_1\left(\frac{d}{dt}\right) V_1\left(-\frac{d}{dt}\right) y(t)$$

при соответствующей системе граничных условий. Общее решение этого уравнения имеет вид

$$y(t) = \sum_{k=1}^{2N} c_k \exp\{t z_k(\delta)\},$$

где N — степень полинома $G_1(i\omega) V_2(i\omega)$; $z_k(\delta)$ — корни характеристического многочлена, соответствующего приведенному дифференциальному уравнению. Константы c_k

находятся из системы линейных алгебраических уравнений, получающихся при удовлетворении граничных условий. Если $D(\delta)$ — детерминант этой системы уравнений и

$$E(\delta) = \frac{D(\delta)}{\prod\limits_{k>j} [z_k(\delta) - z_j(\delta)]},$$

то второе слагаемое в правой части неравенства (3.146) может быть представлено в виде [13]

$$\sum_{k=1}^{\infty} \ln \lambda_k = \frac{E(-1)}{E(0)}. \tag{3.146'}$$

Заметим, что для рассматриваемого случая, когда энергетический спектр стационарного (в широком смысле) нормального случайного процесса представляет дробно-рациональную функцию частоты, сформулировано необходимое и достаточное условие сингулярности [1, 20] при проверке гипотезы H_0 о том, что наблюдаемая реализация принадлежит процессу со спектром $F_0(\omega)$, против альтернативы H_1, что она принадлежит процессу, энергетический спектр которого равен $F_1(\omega)$. Это условие состоит в том, чтобы

$$\lim_{\omega\to\infty} \frac{F_1(\omega)}{F_0(\omega)} \neq 1. \tag{3.147}$$

Указанное выше [см. (3.130)] сингулярное правило *) соответствует частному случаю (3.147), когда $\frac{F_1(\omega)}{F_0(\omega)} = \mu^2 \neq \neq 1$. Если при $\omega \to \infty$ предел отношения дробно-рациональных энергетических спектров равен единице, то всегда будет иметь место регулярный случай, которому соответствуют отличные от нуля вероятности ошибочных решений. Достаточным условием сингулярности является также существование конечного интервала частот, на котором хотя бы один из энергетических спектров, $F_1(\omega)$ или $F_0(\omega)$, тождественно равен нулю. Поэтому использование математической модели случайного процесса с ограниченным энергетическим спектром (см. § 3.2.3) приводит к сингулярности.

---

*) Заметим, однако, что сингулярное правило (3.130) имеет место не только для дробно-рационального, а для произвольного энергетического спектра $F_1(\omega) = \mu^2 F_0(\omega)$.

Наконец, укажем, что регулярный случай будет иметь место всегда, если (при любой из двух гипотез) нормальный процесс содержит аддитивную компоненту в виде *белого шума* одинаковой интенсивности, так как условие (3.147) безошибочной проверки гипотез основывается на использовании различия высокочастотной части энергетического спектра. Так как белый шум (представляющий, например, тепловые шумы) всегда присутствует в любых реальных устройствах, то добавление его устраняет *парадокс сингулярности* и приближает математическую модель к изучаемому физическому процессу (см. также [21]).

3.6. ОЦЕНКИ ХАРАКТЕРИСТИК СЛУЧАЙНОГО ПРОЦЕССА

3.6.1. Оценка параметров распределения случайного процесса по его реализации. В этом разделе будет рассмотрен другой тип статистических выводов относительно неизвестных характеристик распределения случайного процесса по наблюдаемой на конечном интервале времени реализации этого случайного процесса. Речь идет об оценке неизвестных параметров одномерной или многомерной функции распределения случайного процесса. Аналогично тому, как результаты, приведенные в § 3.5, распространяли на случайные процессы теорию проверки статистических гипотез по дискретным выборкам конечного размера из распределения случайных величин, здесь будет дано обобщение теории оценок параметров функций распределения случайных величин (см. § 2.2).

В отличие от оценок параметров функции распределения случайной величины, представляющих функции многих переменных от выборочных значений [см. (2.14)], оценки (статистики) параметров $\vartheta_1, \ldots, \vartheta_m$ функции распределения случайного процесса являются *функционалами* от его реализации $x(t)$:

$$\hat{\vartheta}_i = g_i[x(t)], \quad i = 1, \ldots, m. \tag{3.148}$$

Определение состоятельности и несмещенности оценок (3.148) не отличается от соответствующих определений в § 2.2.2 и 2.2.3 с необходимым добавлением, состоящим в том, что усреднение производится не в конечномерном

пространстве выборок, а в *функциональном* пространстве реализаций *).

Дальнейшие определения и результаты в теории оценок параметров случайных величин связаны с функцией правдоподобия, представляющей плотность распределения точки выборочного пространства [см. (3.13)]. Как уже подчеркивалось в конце § 3.4.1, предел функции правдоподобия, когда размерность пространства выборок неограниченно возрастает, не существует. Это казалось бы создает препятствия для формального обобщения результатов теории оценок, данных во второй главе, на случайные процессы. Для преодоления этого препятствия вместо несуществующего функционала правдоподобия для построения теории оценок параметров случайных процессов вводится **) *функционал отношения правдоподобия*, который в регулярном случае положительный и ограниченный [см. (3.51)].

Так, информация по Фишеру об оцениваемом параметре ϑ, содержащаяся в реализации $x(t)$ случайного процесса, определяется по формуле

$$I(\vartheta) = m_1\left\{\left(\frac{\partial}{\partial\vartheta}\ln l\,[x(t)\,|\,\vartheta]\right)^2\right\}. \qquad (3.149)$$

где символ m_1 означает усреднение по множеству реализаций, а $l\,[x(t)\,|\,\vartheta]$ — функционал отношения правдоподобия при фиксированном значении параметра. Минимум дисперсии оценки определяется обобщенным неравенством Рао — Крамера (см. § 2.2.5)

$$M_2\{\hat{\vartheta}\} \geq \frac{\left[1 + \dfrac{db(\vartheta)}{d\vartheta}\right]^2}{I(\vartheta)}, \qquad (3.150)$$

где *смещение*

$$b(\vartheta) = m_1\{g\,[x(t)]\} - \vartheta. \qquad (3.151)$$

Оценка, для которой в (3.150) достигается равенство, называется *эффективной*.

---

*) Об интегрировании в функциональном пространстве см., например: Гельфанд И. М. и Яглом А. М. Интегрирование в функциональных пространствах. «Успехи математических наук», 1956, вып. 1 (67).

**) Подробнее см. [2 и 4].

Оценка максимального правдоподобия параметра ϑ определяется из уравнения [см. (2.54)]

$$\frac{\partial \ln l\,[x(t)\,|\,\vartheta]}{\partial \vartheta} = 0. \qquad (3.152)$$

Если существует эффективная оценка, то она является оценкой максимального правдоподобия. Обратное утверждение, вообще говоря, неверно.

Аналогичным образом обобщаются и байесовские оценки. Если параметр ϑ случайный и известна его априорная плотность вероятности $w_1(\vartheta)$, то вычисляется апостериорная плотность вероятности параметра ϑ по наблюдаемой реализации $x(t)$ случайного процесса

$$W_1[\vartheta\,|\,x(t)] = \frac{w_1(\vartheta)\, l\,[x(t)\,|\,\vartheta]}{\int\limits_{-\infty}^{\infty} w_1(\vartheta)\, l\,[x(t)\,|\,\vartheta]\, d\vartheta} \qquad (3.153)$$

и далее определяется средний (байесовский) риск

$$R = m_1\left\{\int\limits_{-\infty}^{\infty} \Pi\,(g\,[x(t)], \quad W_1[\vartheta\,|\,x(t)]\, d\vartheta\right\}, \qquad (3.154)$$

где $\Pi(\hat{\vartheta}, \vartheta)$ — функция потерь (см. § 2.3.4).

Обозначая

$$J\,[x(t)\,|\,g] = \int\limits_{-\infty}^{\infty} \Pi\,(g, \vartheta)\, W_1[\vartheta\,|\,x(t)]\, d\vartheta, \qquad (3.155)$$

находим байесовскую оценку $\hat{\vartheta} = g\,[x(t)]$ параметра ϑ из условия

$$\frac{\partial J}{\partial g} = 0. \qquad (3.156)$$

Естественным образом приведенные выше определения и соотношения обобщаются на случай совместных оценок совокупности параметров $\vartheta_1, \ldots, \vartheta_m$. Эти обобщения будут приведены ниже в связи с конкретно формулируемыми задачами.

3.6.2. Оценки максимального правдоподобия параметров детерминированного слагаемого. Предположим, что $x(t)$ — наблюдаемая на интервале $(-T, T)$ реализация процесса, представляющего *сумму* нормального случайного

процесса $\xi(t)$ с нулевым средним и известной корреляционной функцией $B(t, y)$ и детерминированного процесса $s(t; \vartheta_1, \ldots, \vartheta_m)$, зависящего от неизвестных параметров $\vartheta_1, \ldots, \vartheta_m$. Задача состоит в том, чтобы, используя наблюдаемую реализацию $x(t)$, *оценить* неизвестные параметры $\vartheta_1, \ldots, \vartheta_m$ детерминированного слагаемого. Для этого воспользуемся выражением логарифма функционала отношения правдоподобия (3.61), полагая $s_1(t) = s(t; \vartheta_1, \ldots, \vartheta_m)$, $s_0(t) \equiv 0$:

$$\ln l[x(t) \mid \vartheta_1, \ldots, \vartheta_m] = \int_{-T}^{T} V(t; \vartheta_1, \ldots, \vartheta_m)\left[x(t) - \frac{1}{2} s(t; \vartheta_1, \ldots, \vartheta_m)\right] dt, \quad (3.157)$$

где $V(t)$ – решение неоднородного линейного интегрального уравнения

$$\int_{-T}^{T} B(t, u) V(u; \vartheta_1, \ldots, \vartheta_m)\, du = s(t; \vartheta_1, \ldots, \vartheta_m), \quad |t| \leqslant T. \quad (3.158)$$

Принимая максимум функционала отношения правдоподобия как критерий качества оценок $\hat{\vartheta}_1, \ldots, \hat{\vartheta}_m$ неизвестных параметров, находим из (3.157) систему уравнений, которой должны удовлетворять эти оценки:

$$\frac{\partial}{\partial \vartheta_i} \ln [x(t) \mid \vartheta_1, \ldots, \vartheta_m] = 0, \quad i = 1, \ldots, m,$$

или

$$\int_{-T}^{T} \frac{\partial}{\partial \vartheta_i} V(t; \vartheta_1, \ldots, \vartheta_m)\left[x(t) - \frac{1}{2} s(t; \vartheta_1, \ldots, \vartheta_m)\right] dt - \frac{1}{2} \int_{-T}^{T} V(t; \vartheta_1, \ldots, \vartheta_m) \frac{\partial}{\partial \vartheta_i} s(t; \vartheta_1, \ldots, \vartheta_m)\, dt = 0. \quad (3.159)$$

Но из (3.158) следует:

$$\int_{-T}^{T} V(t; \vartheta_1, \ldots, \vartheta_m) \frac{\partial}{\partial \vartheta_i} s(t; \vartheta_1, \ldots, \vartheta_m)\, dt = \int_{-T}^{T} V(t; \vartheta_1, \ldots, \vartheta_m)\left[\int_{-T}^{T} B(t, u) \frac{\partial}{\partial \vartheta_i} V(u; \vartheta_1, \ldots, \vartheta_m)\, du\right] dt =$$

$$= \int_{-T}^{T} \frac{\partial}{\partial \vartheta_i} V(u; \vartheta_1, \ldots, \vartheta_m) \times$$

$$\times \left[\int_{-T}^{T} B(t, u) V(t; \vartheta_1, \ldots, \vartheta_m)\, dt \right] du$$

$$= \int_{-T}^{T} \frac{\partial}{\partial \vartheta_i} V(u; \vartheta_1, \ldots, \vartheta_m)\, s(u; \vartheta_1, \ldots, \vartheta_m)\, du. \qquad (3.160)$$

Подставляя (3.160) в (3.159), получим окончательно

$$\int_{-T}^{T} \frac{\partial}{\partial \vartheta_i} V(t; \vartheta_1, \ldots, \vartheta_m) [x(t) - s(t; \vartheta_1, \ldots, \vartheta_m)]\, dt = 0,$$

$$i = 1, \ldots, m. \qquad (3.161)$$

Решая систему уравнений (3.161) относительно неизвестных $\vartheta_1, \ldots, \vartheta_m$, можно найти условные оценки этих параметров $\hat{\vartheta}_i = g_i\,[x(t)]$, $i = 1, \ldots, m$, удовлетворяющие принятому критерию максимального правдоподобия.

Нетрудно обобщить систему уравнений (3.161) максимального правдоподобия на случай комплексного представления процессов. Используя (3.78), получим

$$\int_{-T}^{T} \frac{\partial}{\partial \vartheta_i} V(t; \vartheta_1, \ldots, \vartheta_m) [\overline{z(t) - s(t; \vartheta_1, \ldots, \vartheta_m)]\, dt} +$$

$$+ \int_{-T}^{T} \frac{\partial}{\partial \vartheta_i} \overline{V(t; \vartheta_1, \ldots, \vartheta_m)} \times$$

$$\times [z(t) - s(t; \vartheta_1, \ldots, \vartheta_m)]\, dt = 0, \qquad (3.161')$$

где $V(t; \vartheta_1, \ldots, \vartheta_m)$ определяется из комплексного интегрального уравнения типа (3.77).

Если процесс $\xi(t)$ представляет белый шум со спектральной плотностью N_0, то из (3.158) следует, что

$$V(t; \vartheta_1, \ldots, \vartheta_m) = \frac{1}{N_0} s(t; \vartheta_1, \ldots, \vartheta_m) \qquad (3.162)$$

и система (3.161) значительно упрощается:

$$\int_{-T}^{T} \frac{\partial}{\partial \vartheta_i} s(t; \vartheta_1, \ldots, \vartheta_m) [x(t) - s(t; \vartheta_1, \ldots, \vartheta_m)]\, dt = 0,$$

$$i = 1, \ldots, m. \qquad (3.163)$$

В частном случае одного неизвестного параметра ϑ из (3.161) при $m=1$ следует:

$$\int_{-T}^{T} \frac{\partial V(t;\ \vartheta)}{\partial \vartheta} [x(t) - s(t;\ \vartheta)]\, dt = 0, \qquad (3.164)$$

а для белого шума соответственно из (3.163) находим

$$\int_{-T}^{T} \frac{\partial s(t;\ \vartheta)}{\partial \vartheta} [x(t) - s(t;\ \vartheta)]\, dt = 0. \qquad (3.165)$$

3.6.3. Оценка амплитуды. Рассмотрим в качестве примера оценку амплитуды детерминированного процесса $as(t)$. Уравнение максимального правдоподобия (3.164) для получения этой оценки преобразуется к виду

$$\int_{-T}^{T} V(t)\, [x(t) - as(t)]\, dt = 0,$$

откуда

$$\hat{a} = \frac{\int_{-T}^{T} V(t)\, x(t)\, dt}{\int_{-T}^{T} V(t)\, s(t)\, dt}, \qquad (3.166)$$

где $V(t)$ в соответствии с (3.158) представляет решение линейного интегрального уравнения

$$\int_{-T}^{T} B(t, u)\, V(u)\, du = s(t), \qquad |t| \leqslant T. \qquad (3.167)$$

Оценку максимального правдоподобия амплитуды детерминированного процесса в аддитивном нормальном белом шуме получаем подстановкой (3.162) при $m=1$ в (3.166):

$$\hat{a} = \frac{\int_{-T}^{T} s(t)\, x(t)\, dt}{\int_{-T}^{T} s^2(t)\, dt}. \qquad (3.168)$$

Заметим, что при $s(t) \equiv 1$ из (3.168) следует

$$\hat{a} = \frac{1}{2T} \int_{-T}^{T} x(t)\, dt, \qquad (3.169)$$

т. е. оценка постоянного среднего a получается усреднением по времени реализации наблюдаемой на интервале $(-T,\ T)$. По аналогии с дискретным случаем оценку (3.169) называют иногда *среднеарифметической*.

Обозначим

$$h(t) = \frac{V(t)}{\int_{-T}^{T} V(t)\, s(t)\, dt}. \qquad (3.170)$$

Тогда оценку (3.166) можно записать в виде

$$\hat{a} = \int_{-T}^{T} h(t)\, x(t)\, dt. \qquad (3.171)$$

Следовательно, оценка максимального правдоподобия амплитуды детерминированного слагаемого получается интегрированием на интервале наблюдения реализации случайного процесса с весом, зависящим от вида детерминированного слагаемого и корреляционной функции процесса. Оценки такого вида называют *линейными*.

Рассмотрим основные свойства оценки (3.171). Прежде всего ясно, что эта оценка несмещенная. Действительно, для любого фиксированного T из (3.170) и (3.171) следует:

$$m_1\{\hat{a}\} = \int_{-T}^{T} h(t)\, m_1\{x(t)\}\, dt = a \int_{-T}^{T} h(t)\, s(t)\, dt = a, \qquad (3.172)$$

так как $m_1\{x(t)\} = as(t)$.

Дисперсия оценки $\hat{a}$ равна

$$M_2\{\hat{a}\} = m_1\{(\hat{a} - a)^2\} = \int_{-T}^{T}\int_{-T}^{T} h(t)\, h(u)\, m_1\{[x(t) - as(t)]\,[x(u) - as(u)]\}\, dt\, du = \int_{-T}^{T}\int_{-T}^{T} h(t)\, h(u)\, B(t,\ u)\, dt\, du \qquad (3.173)$$

или с учетом (3.170) и (3.167)

$$M_2\{\hat{a}\} = \frac{\int\limits_{-T}^{T}\int\limits_{-T}^{T} V(t)\,V(u)\,B(t,u)\,dt\,du}{\left(\int\limits_{-T}^{T} V(t)\,s(t)\,dt\right)^2} = \frac{1}{\int\limits_{-T}^{T} V(t)\,s(t)\,dt}. \quad (3.174)$$

Однако знаменатель в правой части (3.174) совпадает со средним значением квадрата производной по оцениваемому параметру логарифма функционала отношения правдоподобия. Действительно, информация по Фишеру в соответствии с (3.149) равна

$$I(a) = m_1\left\{\left(\frac{\partial}{\partial a}\ln l\,[x(t),\,a]\right)^2\right\} =$$

$$= m_1\left\{\int\limits_{-T}^{T}\int\limits_{-T}^{T} V(t)\,V(u)\,[x(t)-as(t)]\,[x(u)-as(u)]\,dt\,du\right\} =$$

$$\int\limits_{-T}^{T}\int\limits_{-T}^{T} V(t)\,V(u)\,B(t,u)\,dt\,du = \int\limits_{-T}^{T} V(t)\,s(t)\,dt. \quad (3.175)$$

Следовательно, из неравенства Рао — Крамера [см. (3.150)] можно сделать вывод о том, что несмещенная линейная оценка (3.171) имеет минимальную дисперсию в классе всех возможных оценок амплитуды детерминированной составляющей, т. е. является *эффективной оценкой* этого параметра.

Так как оценка (3.171) максимального правдоподобия является линейным функционалом от реализации нормального случайного процесса, она представляет нормальную случайную величину со средним значением a (в силу несмещенности) и дисперсией $M_2\{\hat{a}\}$, определяемой по формуле (3.174). Это дает возможность довольно просто указать *интервальную* оценку для амплитуды [см. (2.52)]

$$P\{a(1-\varepsilon) < \hat{a} < a(1+\varepsilon)\} =$$

$$= \frac{1}{\sqrt{2\pi}}\int\limits_{-\frac{a\varepsilon}{\sqrt{M_2}}}^{\frac{a\varepsilon}{\sqrt{M_2}}} e^{-\frac{u^2}{2}}\,du = 2F\left(\frac{a\varepsilon}{\sqrt{M_2}}\right) - 1 = \gamma, \quad (3.176)$$

откуда следует, что длина доверительного интервала равна [см. (3.174)]

$$2\varepsilon a = 2\left[\int_{-T}^{T} V(t)\, s(t)\, dt\right]^{-\frac{1}{2}} x_{\frac{1-\gamma}{2}}, \qquad (3.176')$$

где $x_{\frac{1-\gamma}{2}}$ — процентная точка нормального распределения, определяемая по коэффициенту доверия γ.

В случае белого шума $N_0 V(t) = s(t)$ и дисперсия (минимальная) оценки $\hat{a}$ определяется по формуле

$$M_2\{\hat{a}\} = \frac{N_0}{\int_{-T}^{T} s^2(t)\, dt}. \qquad (3.177)$$

Следует отметить, что хотя оценка (3.171) представляет оценку максимального правдоподобия амплитуды детерминированного процесса в том случае, когда он комбинируется аддитивно с *нормальным* случайным процессом, эта оценка сохраняет некоторые важные свойства и для процессов $\xi(t)$, *отличных от нормального*. Прежде всего ясно, что эта оценка всегда несмещенная. Кроме того, можно показать, что в *классе линейных оценок* она имеет минимальную дисперсию. Действительно, пусть оценка амплитуды равна

$$\hat{a} = \int_{-T}^{T} g(t)\, x(t)\, dt \qquad (3.178)$$

при условии несмещенности

$$\int_{-T}^{T} g(t)\, s(t)\, dt = 1. \qquad (3.178')$$

Тогда по аналогии с (3.173)

$$M_2\{\hat{a}\} = \int_{-T}^{T}\int_{-T}^{T} g(t)\, g(u)\, B(t, u)\, dt\, du. \qquad (3.179)$$

Найдем нижнюю границу значений дисперсии при всевозможных *) функциях $g(t)$ и соблюдении условия (3.178′),

*) Предполагается, конечно, что класс функций $g(t)$ ограничен не только условием (3.178′), но и требованием ограниченности $M_2\{\hat{a}\}$, т. е. сходимости в среднеквадратическом стохастического интеграла (1.178).

которое с учетом (3.167) может быть переписано в виде

$$\int_{-T}^{T}\int_{-T}^{T} g(t)\,V(u)\,B(t,u)\,dt\,du = 1.$$

Воспользовавшись неравенством Буняковского — Шварца (см., например, сноску на стр. 84 в первой книге), получим

$$1 = \left(\int_{-T}^{T}\int_{-T}^{T} g(t)\,V(t)\,B(t,u)\,dt\,du\right)^2 \leqslant$$

$$\int_{-T}^{T}\int_{-T}^{T} V(t)\,V(u)\,B(t,u)\,dt\,du \int_{-T}^{T}\int_{-T}^{T} g(t)\,g(u)\,B(t,u)\,dt\,du,$$

откуда следует, что [см. (3.179)]

$$M_2\{\hat{a}\} \geqslant \frac{1}{\int_{-T}^{T}\int_{-T}^{T} V(t)\,V(u)\,B(t,u)\,dt\,du}$$

или (см. (3.174))

$$M_2\{\hat{a}\} \geqslant \frac{1}{\int_{-T}^{T} V(t)\,s(t)\,dt}. \qquad (3.180)$$

Таким образом, нижняя граница дисперсий всевозможных линейных оценок амплитуды a детерминированного процесса $as(t)$ в произвольном аддитивном случайном процессе с корреляционной функцией $B(t, u)$ равна $\left(\int_{-T}^{T} V(t)\,s(t)\,dt\right)^{-1}$, где $V(t)$ — решение интегрального уравнения (3.167). Эта граница достигается в том случае, когда $g(t) = h(t)$ [см. (3.170)]. Во всех приведенных выше результатах молчаливо предполагалось, что

$$\int_{-T}^{T} V(t)\,s(t)\,dt = \int_{-T}^{T}\int_{-T}^{T} V(t)\,V(u)\,B(t,u)\,dt\,du > 0, \qquad (3.181)$$

т. е. корреляционная функция является ядром положительно определенной квадратичной формы. Это условие, как отмечалось выше, соответствует регулярному случаю. Если же

$$\int_{-T}^{T} V(t)\,s(t)\,dt = 0,$$

то имеет место сингулярный случай. Наличие аддитивного белого шума является достаточным условием для справедливости предположения (3.181), так как при

$$B(t, u) = B_1(t, u) + N_0\delta(t-u)$$

$$\int_{-T}^{T} V(t)\, s(t)\, dt = \int_{-T}^{T}\int_{-T}^{T} V(t)\, V(u)\, B_1(t, u)\, dt\, du + + N_0 \int_{-T}^{T} V^2(t)\, dt > 0, \qquad (3.181')$$

потому что первое слагаемое в (3.181′) неотрицательное, а второе — положительное.

Таким образом, наличие аддитивного белого шума сразу же исключает возможность появления сингулярности.

3.6.4. Байесовские оценки. Используя обобщение методики, изложенной в § 2.4.5, нетрудно в принципе определить и байесовские оценки случайных параметров $\vartheta_1, \ldots, \vartheta_m$ квазидетерминированного слагаемого $s(t; \vartheta_1, \ldots, \vartheta_m)$ по наблюдаемой реализации $x(t)$.

Байесовские оценки $\hat{\vartheta}_1 = g_1[x(t)], \ldots, \hat{\vartheta}_m = g_m[x(t)]$ неизвестных параметров $\vartheta_1, \ldots, \vartheta_m$, минимизирующие средний риск

$$R = m_1\{J[x(t) \mid g_1, \ldots, g_m]\}, \qquad (3.182)$$

находятся из системы уравнений [см. (2.118)]

$$\frac{\partial J}{\partial g_i} = 0, \qquad i = 1, \ldots, m, \qquad (3.183)$$

где

$$J[x(t) \mid g_1, \ldots, g_m] = = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \Pi(\hat{\vartheta}_1, \ldots, \hat{\vartheta}_m, \vartheta_1, \ldots, \vartheta_m) \times \times W_m[\vartheta_1, \ldots, \vartheta_m \mid x(t)]\, d\vartheta_1 \ldots d\vartheta_m \qquad (3.184)$$

и

$$W_m[\vartheta_1, \ldots, \vartheta_m \mid x(t)] = = \frac{w_m(\vartheta_1, \ldots, \vartheta_m)\, l[x(t) \mid \vartheta_1, \ldots, \vartheta_m]}{\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} w_m(\vartheta_1, \ldots, \vartheta_m)\, l[x(t) \mid \vartheta_1, \ldots, \vartheta_m]\, d\vartheta_1 \ldots d\vartheta_m}. \qquad (3.185)$$

Для квадратичной функции потерь получим [см. (2.119)]

$$\hat{\vartheta}_i = \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} \vartheta_i W_m [\vartheta_1, \ldots, \vartheta_m \,|\, x(t)]\, d\vartheta_1 \ldots d\vartheta_m. \qquad (3.186)$$

При ограничениях, указанных во второй главе, эта же байесовская оценка имеет место и для иных (симметричных) функций потерь.

Возвращаясь к рассмотренному выше примеру, найдем байесовскую оценку амплитуды a квазидетерминированного процесса $as(t)$, предположив сначала, что функция потерь П — квадратичная и что амплитуда представляет нормальную случайную величину с точно известными средним a_0 и дисперсией σ_0^2.

Тогда в соответствии с (3.157) и (3.185) апостериорная плотность параметра a, когда имеется отрезок реализации $x(t)$ на интервале $(-T, T)$, равна

$$W_1 [a \,|\, x(t)] = \frac{w_1(a)\, l\,[x(t) \,|\, a]}{\int\limits_{-\infty}^{\infty} w_1(a)\, l\,[x(t) \,|\, a]\, da} =$$

$$= \frac{1}{\sigma_0 \sqrt{2\pi}} e^{-\frac{(a-a_0)^2}{2\sigma_0^2}} \exp\left\{a \int\limits_{-T}^{T} V(t) \left[x(t) - \frac{a}{2} s(t)\right] dt\right\} \times$$

$$\times \left\{\frac{1}{\sigma_0 \sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{(a-a_0)^2}{2\sigma_0^2}} \exp\left[a \int\limits_{-T}^{T} V(t) \times\right.\right.$$

$$\left.\left.\times \left[x(t) - \frac{a}{2} s(t)\right] dt\right] da\right\}^{-1},$$

где $V(t)$ представляет решение интегрального уравнения (3.167). Вычисляя интеграл, заключенный в фигурных скобках, получим после несложных преобразований

$$W_1 [a \,|\, x(t)] = \frac{(1 + \sigma_0^2 s_T)^{1/2}}{\sigma_0 \sqrt{2\pi}} \times$$

$$\times \exp\left\{-\frac{1 + \sigma_0^2 s_T}{2\sigma_0^2} \left(a - \frac{a_0 + \sigma_0^2 x_T}{1 + \sigma_0^2 s_T}\right)^2\right\}, \qquad (3.187)$$

где

$$s_T = \int_{-T}^{T} V(t)\, s(t)\, dt, \qquad (3.188)$$

$$x_T = \int_{-T}^{T} V(t)\, x(t)\, dt. \qquad (3.189)$$

При квадратичной функции потерь *) минимум $J\,[x(t)\,|\,g]$ достигается для оценки, совпадающей с условным средним оцениваемого параметра a (см. § 2.3.6). Тогда из (3.187) непосредственно следует, что байесовская оценка амплитуды квазидетерминированного слагаемого имеет вид

$$\hat{a} = \frac{a_0 + \sigma_0^2 \int_{-T}^{T} V(t)\, x(t)\, dt}{1 + \sigma_0^2 \int_{-T}^{T} V(t)\, s(t)\, dt} = \frac{a_0 + \sigma_0^2 s_T \dfrac{x_T}{s_T}}{1 + \sigma_0^2 s_T} \qquad (3.190)$$

или

$$\hat{a} = a_0 + \frac{\sigma_0^2}{1 + \sigma_0^2 s_T} \int_{-T}^{T} V(t)\,[x(t) - a_0 s(t)]\, dt =$$
$$= a_0 + \frac{\sigma_0^2}{1 + \sigma_0^2 s_T}(x_T - a_0 s_T). \qquad (3.190')$$

Сравнивая (3.190) с (3.166), нетрудно прийти к выводу, что байесовская оценка представляет среднее взвешенное двух величин: оценки максимального правдоподобия и априорного среднего a_0, причем отношение веса, приписываемого первой величине, к весу второй равно

$$\mu = \sigma_0^2 s_T = \frac{\sigma_0^2}{\left(\int_{-T}^{T} V(t)\, s(t)\, dt\right)^{-1}}, \qquad (3.191)$$

т. е. отношению дисперсии априорного распределения параметра к дисперсии оценки максимального правдоподобия.

Если указанное отношение неограниченно возрастает (т. е. дисперсия оценки максимального правдоподобия много меньше дисперсии априорного распределения), то

$$\hat{a} \sim \frac{x_T}{s_T}, \qquad (3.191')$$

*) Из § 2.3.8 следует, что приводимые ниже результаты справедливы также и при любой симметричной функции потерь.

т. е. байесовская оценка приближается к оценке максимального правдоподобия. Если же, напротив, дисперсия априорного распределения много меньше дисперсии оценки максимального правдоподобия, то

$$\hat{a} \sim a_0, \tag{3.192}$$

т. е. наблюдаемая реализация не влияет на оценку, которая принимается равной априорному среднему оцениваемого параметра.

Когда случайная составляющая процесса представляет белый шум, то согласно (3.167) $N_0 V(t) = s(t)$, и тогда из (3.190) следует:

$$\hat{a} = \frac{a_0 + \frac{\sigma_0^2}{N_0} \int\limits_{-T}^{T} s(t)\, x(t)\, dt}{1 + \frac{\sigma_0^2}{N_0} \int\limits_{-T}^{T} s^2(t)\, dt} \tag{3.192'}$$

или

$$\hat{a} = \frac{a_0 + \mu \hat{a}_{\text{мп}}}{1 + \mu}, \tag{3.192''}$$

где $\hat{a}_{\text{мп}}$ — оценка максимального правдоподобия (3.168). Величина

$$\mu = \frac{\sigma_0^2}{N_0} \int\limits_{-T}^{T} s^2(t)\, dt,$$

т. е. равна отношению дисперсии априорного распределения к дисперсии оценки максимального правдоподобия [см. (3.177)].

Рассмотрим некоторые свойства байесовской оценки (3.190) амплитуды квазидетерминированного слагаемого. Прежде всего видно, что эта оценка получается путем *линейной* обработки наблюдаемой реализации $x(t)$ и, следовательно, так же как и исходный процесс, имеет нормальное распределение (как и оценка максимального правдоподобия). Найдем среднее и дисперсию байесовской оценки $\hat{a}$, причем сначала получим условные средние и дисперсию при фиксированном a, а затем безусловные, усредняя по a.

Из (3.190′) следует:

$$m_1\{\hat{a}\,|\,a\} = a_0 + (a - a_0)\,\frac{\sigma_0^2 s_T}{1+\sigma_0^2 s_T}\,,$$

$$M_2\{\hat{a}\,|\,a\} = \frac{\sigma_0^4 s_T}{(1+\sigma_0^2 s_T)^2}\,[1 + (a-a_0)^2 s_T],$$

и так как $m_1\{a\} = a_0$, $m_1\{(a-a_0)^2\} = \sigma_0^2$, то, усредняя по a, получим

$$m_1\{\hat{a}\} = a_0, \tag{3.193}$$

$$M_2\{\hat{a}\} = \frac{\sigma_0^4 s_T}{1+\sigma_0^2 s_T} = \frac{\mu}{1+\mu}\,\sigma_0^2. \tag{3.193′}$$

При $\mu \to 0$ дисперсия байесовской оценки также стремится к нулю, а при $\mu \to \infty$ асимптотическое значение этой дисперсии равно σ_0^2.

Заметим, что функция $J\,[x(t)\,|\,\hat{a}]$ в рассматриваемом примере в соответствии с (3.189), (3.187) и (3.190) равна

$$J\,[x(t)\,|\,\hat{a}] = \int\limits_{-\infty}^{\infty} (a-\hat{a})^2\, W_1\,[a\,|\,x(t)]\,da =$$

$$= \frac{1}{\sqrt{2\pi}}\left(\frac{\sigma_0^2}{1+\sigma_0^2 s_T}\right)^{-\frac{1}{2}} \int\limits_{-\infty}^{\infty} (a-\hat{a})^2 \times$$

$$\times \exp\left[-\frac{1+\sigma_0^2 s_T}{\sigma_0^2}(a-\hat{a})^2\right] da = \frac{\sigma_0^2}{1+\sigma_0^2 s_T}\,, \tag{3.194}$$

т. е. *не зависит* от $x(t)$. Отсюда следует, что средний риск равен

$$R = m_1\{J\,[x(t)\,|\,\hat{a}]\} = \frac{\sigma_0^2}{1+\sigma_0^2 s_T} = \frac{\mu}{1+\mu}\cdot\frac{1}{s_T}\,. \tag{3.195}$$

При $\sigma_0 \to 0$ величина среднего риска также стремится к нулю, так как в этом случае априорное распределение оцениваемого параметра приближается к дельта-функции.

При $\sigma_0 \to \infty$

$$R \to R_{max} = \frac{1}{s_T}\,, \tag{3.196}$$

т. е. к величине дисперсии оценки максимального правдоподобия.

Таким образом, априорное распределение, для которого дисперсия много больше дисперсии оценки максимального правдоподобия, является *наименее благоприятным* и, сле-

довательно, оценка (3.166) максимального правдоподобия является также и *минимаксной*.

3.6.5. Асимптотические свойства байесовской оценки. При произвольном априорном распределении $w_1(a)$ амплитуды квазидетерминированного сигнала байесовская оценка этой амплитуды при квадратичной функции потерь равна в соответствии с (3.186)

$$\hat{a} = \int_{-\infty}^{\infty} a W_1 [a | x(t)]\, da, \qquad (3.197)$$

где $x(t)$ — наблюдаемая на интервале $(-T,\ T)$ реализация суммы сигнала $as(t)$ и нормального случайного процесса с нулевым средним и известной корреляционной функцией $B(t,\ y)$.

Подставляя в (3.197) выражение для апостериорной плотности оцениваемого параметра [см. (3.187)], используя выражение для функционала отношения правдоподобия и обозначения (3.188), (3.189), запишем (3.197) в виде

$$\hat{a} \quad \frac{\int\limits_{-\infty}^{\infty} a w_1(a) \exp\left(a x_T - \frac{a^2}{2} s_T\right) da}{\int\limits_{-\infty}^{\infty} w_1(a) \exp\left(a x_T - \frac{a^2}{2} s_T\right) da}.$$

Дополняя экспоненты в подынтегральных выражениях до полных квадратов, после элементарных алгебраических преобразований находим

$$\hat{a} = \frac{\int\limits_{-\infty}^{\infty} a w_1(a) \sqrt{\frac{s_T}{2\pi}} \exp\left[\frac{s_T}{2}\left(a - \frac{x_T}{s_T}\right)^2\right] da}{\int\limits_{-\infty}^{\infty} w_1(a) \sqrt{\frac{s_T}{2\pi}} \exp\left[\frac{s_T}{2}\left(a - \frac{x_T}{s_T}\right)^2\right] da}. \qquad (3.198)$$

Когда дисперсия оценки максимального правдоподобия $s_T^{-1} \to 0$, функция

$$\sqrt{\frac{s_T}{2\pi}} \exp\left[\frac{s_T}{2}\left(a - \frac{x_T}{s_T}\right)^2\right] \to \delta\left(a - \frac{x_T}{s_T}\right)$$

(см. стр. 69 в первой книге), и из (3.198), предполагая непрерывность $w_1(a)$ и используя фильтрующее свойство дель-

та-функции, получаем асимптотическую формулу для байесовской оценки амплитуды при $s_T \to \infty$:

$$\hat{a} \sim \frac{x_T}{s_T} = \frac{\int\limits_{-T}^{T} V(t)\, x(t)\, dt}{\int\limits_{-T}^{T} V(t)\, s(t)\, dt}. \qquad (3.199)$$

Таким образом, независимо от вида априорного распределения $w_1(a)$ амплитуды a байесовская оценка $\hat{a}$ приближается при $s_T \to \infty$ к оценке максимального правдоподобия (3.166).

В том случае, когда аддитивный случайный процесс представляет белый шум,

$$s_T = \frac{1}{N_0} \int\limits_{-T}^{T} s^2(t)\, dt,$$

т. е. указанная асимптотика имеет место при неограниченном увеличении отношения энергии детерминированного процесса на интервале наблюдения к спектральной плотности аддитивного белого шума.

Так как несмещенная оценка максимального правдоподобия (3.166) имеет нормальное распределение и ее дисперсия равна s_T^{-1}, то асимптотическое (при $s_T \to \infty$) условное распределение байесовской оценки может быть записано в виде

$$W_1(\hat{a} \mid a) \sim \sqrt{\frac{s_T}{2\pi}} \exp\left[-\frac{s_T}{2}(\hat{a} - a)^2\right], \qquad (3.200)$$

откуда следует, что безусловное распределение байесовской оценки $\hat{a}$ при $s_T \to \infty$ асимптотически приближается к

$$W_1(\hat{a}) = \int\limits_{-\infty}^{\infty} w_1(a)\, W_1(\hat{a} \mid a)\, da \sim w_1(\hat{a}), \qquad (3.200')$$

т. е. к априорному распределению неизвестного параметра, в котором переменная a заменяется оценкой максимального правдоподобия этого параметра.

3.6.6. Обобщение результатов. Вернемся к общему случаю, когда детерминированное слагаемое $s(t; \vartheta_1, \ldots, \vartheta_m)$ содержит произвольное конечное число неизвестных параметров. Случайное слагаемое, как и прежде, представляет нормальный случайный процесс с нулевым средним и изве-

стной корреляционной функцией. Предположим, что функция $s(t;\ \vartheta_1, \ldots, \vartheta_m)$ может быть представлена в виде

$$s(t;\ \vartheta_1, \ldots, \vartheta_m) = \sum_{j=1}^{m} \vartheta_j s_j(t), \tag{3.201}$$

где $s_j(t)$, $j = 1, \ldots, m$ — известные функции.

Найдем совместные оценки максимального правдоподобия параметров $\vartheta_1, \ldots, \vartheta_m$. Подставляя (3.201) в правую часть (3.158), заменим интегральное уравнение (3.158) системой уравнений

$$\int_{-T}^{T} B(t, u)\, V_i(u)\, du = s_i(t), \qquad |t| \leqslant T,$$
$$i = 1, \ldots, m, \tag{3.202}$$

причем функция $V(t;\ \vartheta_1, \ldots, \vartheta_m)$, от которой зависит логарифм функционала отношения правдоподобия (3.157), равна

$$V(t;\ \vartheta_1, \ldots, \vartheta_m) = \sum_{j=1}^{m} \vartheta_j V_j(t). \tag{3.203}$$

Из (3.203) находим

$$\frac{\partial}{\partial \vartheta_i} V(t;\ \vartheta_1, \ldots, \vartheta_m) = V_i(t) \tag{3.204}$$

и, подставляя (3.204) в (3.161), получаем систему уравнений максимального правдоподобия

$$\int_{-T}^{T} V_i(t) \left[x(t) - \sum_{j=1}^{m} \vartheta_j s_j(t) \right] dt = 0, \quad i = 1, \ldots, m, \tag{3.205}$$

или

$$\sum_{j=1}^{m} \vartheta_j \int_{-T}^{T} V_i(t)\, s_j(t)\, dt = \int_{-T}^{T} V_i(t)\, x(t)\, dt,$$
$$i = 1, \ldots, m. \tag{3.205$'$}$$

По аналогии с (3.188) и (3.189) обозначим

$$s_{Tij} = \int_{-T}^{T} V_i(t)\, s_j(t)\, dt, \tag{3.206}$$

$$x_{Ti} = \int_{-T}^{T} V_i(t)\, x(t)\, dt \tag{3.206$'$}$$

и представим систему линейных уравнений относительно (3.205) в виде

$$\sum_{j=1}^{m} s_{Tij}\vartheta_j = x_{Ti}, \qquad i = 1, \ldots, m, \tag{3.207}$$

или в матричной форме

$$\mathbf{s}_T \boldsymbol{\vartheta} = \mathbf{X}_T, \tag{3.207$'$}$$

где $\mathbf{s}_T$ — матрица размером $m \times m$, элементы которой равны s_{Tij}, а $\boldsymbol{\vartheta}$ и $\mathbf{X}_T$ — векторы-столбцы, элементы которых ϑ_j и x_{Ti} соответственно.

Полагая, что для всех j

$$\int_{-T}^{T} s_j^2(t)\,dt < \infty$$

и что $B(t, u)$ — положительно определенная функция (см. § 3.3.1), приходим к заключению, что существует матрица $\mathbf{s}_T^{-1}$, обратная матрице $\mathbf{s}_T$. Тогда решение уравнения (3.207′) приводит к следующим оценкам максимального правдоподобия неизвестных параметров:

$$\hat{\boldsymbol{\vartheta}} = \mathbf{s}_T^{-1}\mathbf{X}_T. \tag{3.208}$$

Если случайная составляющая наблюдаемого процесса представляет белый шум с интенсивностью N_0, то из (3.202) следует, что

$$N_0 V_i(t) = s_i(t), \qquad i = 1, \ldots, m,$$

и, следовательно, матрицы $\mathbf{s}_T$ и $\mathbf{X}_T$ в (3.207′) преобразуются к виду (множитель N_0 сокращается)

$$\mathbf{s}_T = \left\| \int_{-T}^{T} s_i(t)\, s_j(t)\,dt \right\|, \qquad i = 1, \ldots, m, \quad j = 1, \ldots, m, \tag{3.209}$$

$$\mathbf{X}_T = \left\| \int_{-T}^{T} s_i(t)\, x(t)\,dt \right\|, \qquad i = 1, \ldots, m. \tag{3.209$'$}$$

Векторная оценка (3.208) несмещенная. Действительно, из (3.208) находим *)

$$m_1\{\hat{\boldsymbol{\vartheta}}\} = \mathbf{s}_T^{-1} m_1\{\mathbf{X}_T\}. \tag{3.210}$$

*) О среднем значении матриц см. § 2.7.1.

Но

$$m_1\{x_{Ti}\} = \sum_{j=1}^{m} \vartheta_j \int_{-T}^{T} s_j(t) V_i(t)\, dt = \sum_{j=1}^{m} \vartheta_j s_{Tij},$$

и, следовательно,

$$m_1\{\mathbf{X}_T\} = \mathbf{s}_T \boldsymbol{\vartheta}. \qquad (3.210')$$

Подставляя (3.210′) в (3.210), получим

$$m_1\{\hat{\boldsymbol{\vartheta}}\} = \boldsymbol{\vartheta}, \qquad (3.211)$$

что и доказывает утверждение о несмещенности оценки (3.208).

Найдем выражение корреляционной матрицы **M** оценок (3.208) максимального правдоподобия [см. (2.99)]

$$\mathbf{M} = m_1\{(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta})(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta})'\}, \qquad (3.212)$$

где знак штрих указывает на транспонированный вектор (в дальнейшем — на транспонированную матрицу). После несложных преобразований и подстановки вместо $\hat{\boldsymbol{\vartheta}}$ ее выражения из (3.208), получим с учетом симметричности матрицы $\mathbf{s}_T$:

$$\mathbf{M} = m_1\{\hat{\boldsymbol{\vartheta}}\hat{\boldsymbol{\vartheta}}'\} - \boldsymbol{\vartheta}\boldsymbol{\vartheta}' = m_1\{\mathbf{s}_T^{-1}\mathbf{X}_T\mathbf{X}_T'(\mathbf{s}_T^{-1})'\} - \boldsymbol{\vartheta}\boldsymbol{\vartheta}' =$$

$$= \mathbf{s}_T^{-1} m_1\{\mathbf{X}_T\mathbf{X}_T'\}\mathbf{s}_T^{-1} - \boldsymbol{\vartheta}\boldsymbol{\vartheta}'. \qquad (3.213)$$

Но

$$m_1\{x_{Ti}x_{Tj}\} = \int_{-T}^{T}\int_{-T}^{T} B(u, v) V_i(u) V_j(v)\, du\, dv +$$

$$+ \sum_{k=1}^{m} \vartheta_k s_{Tik} \sum_{n=1}^{m} \vartheta_n s_{Tjn},$$

и, следовательно [см. (3.202)],

$$m_1\{\mathbf{X}_T\mathbf{X}_T'\} = \mathbf{s}_T + \mathbf{s}_T\boldsymbol{\vartheta}\boldsymbol{\vartheta}'\mathbf{s}_T. \qquad (3.214)$$

Подставляя (3.214) в (3.213), находим

$$\mathbf{M} = \mathbf{s}_T^{-1}, \qquad (3.215)$$

т. е. корреляционная матрица оценок максимального правдоподобия детерминированного слагаемого (3.201) совпадает с матрицей, обратной матрице $\mathbf{s}_T$ [см. (3.206), а для белого шума (3.209)].

Найдем также элементы информационной матрицы Фишера [см. (2.98)]. Используя левую часть (3.205), получим

$$I^{(i,\,j)} = m_1\left\{\frac{\partial \ln l}{\partial \vartheta_i}\frac{\partial \ln l}{\partial \vartheta_j}\right\} = m_1\left\{\int\limits_{-T}^{T} V_i(u)\times\right.$$

$$\left.\times\left[x(u)-\sum_{k=1}^{m}\vartheta_k s_k(u)\right]du\int\limits_{-T}^{T} V_j(v)\left[x(v)-\sum_{n=1}^{m}\vartheta_n s_n(v)\right]dv\right\} =$$

$$= \int\limits_{-T}^{T}\int\limits_{-T}^{T} B(u,\,v)\,V_i(u)\,V_j(v)\,du\,dv =$$

$$\int\limits_{-T}^{T} V_i(u)\,s_j(u)\,du,$$

что полностью совпадает с (3.206).

Таким образом, информационная матрица совпадает с матрицей $\mathbf{s}_T$, т. е.

$$\mathbf{I} = \mathbf{s}_T, \tag{3.216}$$

и из (3.215) находим, что

$$\mathbf{I}^{-1} = \mathbf{s}_T^{-1} = \mathbf{M}. \tag{3.217}$$

Следовательно [см. (2.99)], оценки максимального правдоподобия (3.208) являются *совместно эффективными*.

Результаты, аналогичные приведенным выше, могут быть получены и в том случае, когда результаты наблюдений представлены не реализацией $x(t)$, а дискретной выборкой (см. задачу 3.6).

Возможны также обобщения результатов § 2.6.4 и на случай совместных байесовских оценок случайных параметров $\vartheta_1, \ldots, \vartheta_m$ квазидетерминированного слагаемого вида (3.201) (см. задачу 3.7, а также [11]).

3.6.7. Оценка параметров корреляционной функции. Рассмотрим задачу оценки неизвестных параметров $\vartheta_1, \ldots, \vartheta_m$ корреляционной функции $B(t, y; \vartheta_1, \ldots, \vartheta_m)$ нормального случайного процесса с нулевым средним. Для того чтобы исключить сингулярность, предположим, что к процессу добавляется белый нормальный шум со спектральной плотностью N_0. Таким образом, корреляционная функция изучаемого случайного процесса $\xi(t)$ представляется суммой

$$B_\xi(t,\,y) = N_0\delta(y-t) + B(t,\,y;\,\vartheta_1\,\ldots,\,\vartheta_m). \tag{3.218}$$

Воспользуемся выражением (3.123) логарифма функционала отношения правдоподобия, полагая в этом выражении

$$B_1(t,\, u) = B_{\xi}(t,\, u) \quad \text{и} \quad B_0(t,\, u) = N_0\delta(t-u)$$

$$\ln l\,[x(t);\, \vartheta_1,\, \ldots,\, \vartheta_m] = \frac{1}{2}\sum_{k=1}^{\infty}[\lambda_k(\vartheta_1,\, \ldots,\, \vartheta_m) - 1] \times$$
$$\times \left[\int_{-T}^{T} x(t)\,\varphi_k(t;\, \vartheta_1,\, \ldots,\, \vartheta_m)\,dt\right]^2 -$$
$$-\frac{1}{2}\sum_{k=1}^{\infty}\ln\lambda_k(\vartheta_1,\, \ldots,\, \vartheta_m), \qquad (3.219)$$

где в соответствии с (3.120) φ_k и λ_k — собственные функции и собственные числа линейного интегрального уравнения

$$\int_{-T}^{T}[B(t,u;\vartheta_1,\, \ldots,\, \vartheta_m) + N_0\delta(t-u) - \lambda N_0\delta(t-u)]\,\varphi(u)\,du = 0$$

или

$$N_0(\lambda - 1)\,\varphi(t) = \int_{-T}^{T} B(t,\, u;\, \vartheta_1,\, \ldots,\, \vartheta_m)\,\varphi(u)\,du, \qquad |t| \leqslant T, \qquad (3.220)$$

причем собственные функции нормированы согласно (3.120′):

$$N_0\lambda_k(\vartheta_1,\, \ldots,\, \vartheta_m)\int_{-T}^{T}\varphi_k(t;\, \vartheta_1,\, \ldots,\, \vartheta_m)\,\varphi_n(t;\, \vartheta_1,\, \ldots,\, \vartheta_m)\,dt =$$
$$= \begin{cases} 1, & k = n, \\ 0, & k \neq n. \end{cases} \qquad (3.221)$$

Принимая максимум функционала отношения правдоподобия как критерий качества оценок $\hat{\vartheta}_1,\, \ldots,\, \hat{\vartheta}_m$ неизвестных параметров, находим, что эти оценки должны удовлетворять системе уравнений

$$\frac{\partial}{\partial\vartheta_i}\ln l\,[x(t);\, \vartheta_1,\, \ldots,\, \vartheta_m] = 0, \qquad i = 1,\, \ldots,\, m, \qquad (3.222)$$

в которую следует подставить выражение $\ln l\,[x(t);\, \vartheta_1,\, \ldots,\, \vartheta_m]$ из (3.219).

Рассмотрим подробно тот случай, когда корреляционная функция $B(t, y)$ известна с точностью до постоянного множителя, т. е.

$$B(t, y) = cR(t, y). \tag{3.223}$$

Для стационарного нормального случайного процесса неизвестным параметром c может быть дисперсия; тогда $R(t, y) = R(y - t)$ представляет известный коэффициент корреляции.

В рассматриваемом случае интегральное уравнение (3.220) преобразуется к виду

$$\varphi(t) = \mu \int_{-T}^{T} R(t, u)\,\varphi(u)\,du, \quad |t| \leqslant T, \tag{3.224}$$

где

$$\mu = \frac{c}{N_0(\lambda - 1)}. \tag{3.225}$$

Из последней формулы также следует, что

$$\lambda = 1 + \frac{c}{N_0\mu}. \tag{3.226}$$

Запишем теперь функционал логарифма отношения правдоподобия, подставляя (3.226) в (3.219). С учетом нормировки $\varphi_k(t; c) = \frac{1}{\sqrt{N_0\lambda_k}}\,\varphi_k(t)$ получим

$$\ln[x(t) \mid c] = \frac{1}{2N_0} \sum_{k=1}^{\infty} \frac{\frac{c}{N_0\mu_k}}{1 + \frac{c}{N_0\mu_k}} \left[\int_{-T}^{T} x(t)\,\varphi_k(t)\,dt\right]^2 - \frac{1}{2} \sum_{k=1}^{\infty} \ln\left(1 + \frac{c}{N_0\mu_k}\right). \tag{3.227}$$

Дифференцируя (3.227) по c, находим уравнение максимального правдоподобия

$$\frac{\partial}{\partial c} \ln l\,[x(t \mid c)] = \frac{1}{2N_0} \sum_{k=1}^{\infty} \left[\frac{1}{N_0\mu_k\left(1 + \frac{c}{N_0\mu_k}\right)} - \frac{c}{N_0^2\mu_k^2\left(1 + \frac{c}{N_0\mu_k}\right)^2}\right] \times \left[\int_{-T}^{T} x(t)\,\varphi_k(t)\,dt\right]^2 - \frac{1}{2} \sum_{k=1}^{\infty} \frac{1}{N_0\mu_k\left(1 + \frac{c}{N_0\mu_k}\right)} = 0$$

или

$$\frac{1}{N_0}\sum_{k=1}^{\infty}\frac{1}{\left(1+\frac{c}{N_0\mu_k}\right)^2\mu_k}\left[\int_{-T}^{T}x(t)\varphi_k(t)\,dt\right]^2=$$

$$=\sum_{k=1}^{\infty}\frac{1}{\left(1+\frac{c}{N_0\mu_k}\right)\mu_k}. \qquad (3.228)$$

Решая уравнение (3.228) относительно c, получаем оценку максимального правдоподобия. В этом уравнении μ_k и $\varphi_k(t)$ — собственные числа и функции интегрального уравнения (3.224). Конечно, вычисление оценки максимального правдоподобия в явном виде из (3.228) весьма трудоемкое. Ясно, что эта оценка получается в результате *нелинейных* операций над наблюдаемой реализацией $x(t)$.

Определим нижнюю границу дисперсий несмещенных оценок параметра c корреляционной функции. Для этого вычислим величину информации по Фишеру [см. (3.149)]:

$$I(c)=m_1\left\{\left(\frac{\partial}{\partial c}\ln l[x(t)\mid c]\right)^2\right\}=$$

$$=\frac{1}{4N_0^2}\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}m_1\left\{\left[\frac{\eta_k^2}{N_0\mu_k\left(1+\frac{c}{N_0\mu_k}\right)^2}-\frac{1}{\mu_k\left(1+\frac{c}{N_0\mu_k}\right)}\right]\times\right.$$

$$\left.\times\left[\frac{\eta_n^2}{N_0\mu_n\left(1+\frac{c}{N_0\mu_n}\right)^2}-\frac{1}{\mu_n\left(1+\frac{c}{N_0\mu_n}\right)}\right]\right\}=$$

$$=\frac{1}{4N_0^2}\sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{\mu_k\mu_n\left(1+\frac{c}{N_0\mu_k}\right)^2\left(1+\frac{c}{N_0\mu_n}\right)^2}\times$$

$$\times m_1\left\{\left(\eta_k^2-N_0-\frac{c}{\mu_k}\right)\left(\eta_n^2-N_0-\frac{c}{\mu_n}\right)\right\},$$

где

$$\eta_k=\int_{-T}^{T}x(t)\varphi_k(t)\,dt$$

— нормальная случайная величина с нулевым средним и дисперсией $N_0+\frac{c}{\mu_k}$, а η_k и η_n при $k\neq n$ независимы. Поэтому при $k\neq n$ $m_1\left\{\left(\eta_k^2-N_0-\frac{c}{\mu_k}\right)\left(\eta_n^2-N_0-\frac{c}{\mu_n}\right)\right\}=0$,

а при $k=n$

$$m_1\left\{\left(\eta_k^2-N_0-\frac{c}{\mu_k}\right)^2\right\}=m_1\{\eta_k^4\}-\left(N_0+\frac{c}{\mu_k}\right)^2=$$
$$=2\left(N_0+\frac{c}{\mu_k}\right)^2.$$

Тогда

$$I(c)=\frac{1}{2N_0^2}\sum_{k=1}^{\infty}{}'\frac{1}{\mu_k^2}\frac{1}{\left(1+\frac{c}{N_0\mu_k}\right)^2}. \qquad (3.229)$$

Согласно неравенству Рао — Крамера [см. (3.150)] находим, что дисперсия несмещенной оценки

$$M_2\{\hat{c}\}\geqslant\frac{2N_0^2}{\sum\limits_{k=1}^{\infty}{}'\frac{1}{\mu_k^2\left(1+\frac{c}{N_0\mu_k}\right)^2}}. \qquad (3.230)$$

Нижнюю границу дисперсии [правую часть неравенства (3.230)] можно с учетом (3.225) переписать в виде

$$M_{2min}=\frac{2N_0^2}{\sum\limits_{k=1}^{\infty}{}'(\mu_k\lambda_k)^{-2}}. \qquad (3.231)$$

Вводя функцию $h(t, y)$ согласно (3.132), которая представляет решение интегрального уравнения [см. (3.134)]

$$\int_{-T}^{T}R(t, y)h(v, y)\,dy+\frac{N_0}{c}h(v, t)=R(t, v), \quad |t|\leqslant T, \qquad (3.232)$$

получаем из (3.231) выражение нижней границы дисперсии оценки параметра c в виде

$$M_{2min}=\frac{2c^2}{\int\limits_{-T}^{T}\int\limits_{-T}^{T}h^2(t, y)\,dt\,dy}, \qquad (3.233)$$

так как из (3.132) следует, что

$$\int_{-T}^{T}\int_{-T}^{T}h^2(t, y)\,dt\,dy=\sum_{k=1}^{\infty}\left(\frac{\lambda_k-1}{\lambda_k}\right)^2=\frac{c^2}{N_0^2}\sum_{k=1}^{\infty}{}'(\mu_k\lambda_k)^{-2}.$$

Если $\frac{N_0}{cT} \gg 1$, то из (3.232) в качестве приближенного решения интегрального уравнения можно принять

$$h(v, t) \approx \frac{c}{N_0} R(t, v). \qquad (3.234)$$

Тогда для нижней границы дисперсий оценки параметра c имеет место асимптотическое соотношение

$$\frac{M_{2min}}{c^2} \sim \frac{2\left(\frac{N_0}{cT}\right)^2}{\frac{1}{T^2}\int\limits_{-T}^{T}\int\limits_{-T}^{T} R^2(u, v)\,du\,dv}. \qquad (3.235)$$

Для стационарного в широком смысле процесса

$$\int\limits_{-T}^{T}\int\limits_{-T}^{T} R^2(u, v)\,du\,dv = \int\limits_{-2T}^{2T} (2T - |\tau|) R^2(\tau)\,d\tau$$

и соотношение (3.235) может быть переписано в виде

$$\frac{M_{2min}}{c^2} \sim \left(\frac{N_0}{cT}\right)^2 k_T, \qquad (3.235')$$

где

$$\frac{1}{k_T} = \frac{1}{T}\int\limits_{-2T}^{2T}\left(1 - \frac{|\tau|}{2T}\right) R^2(\tau)\,d\tau. \qquad (3.236)$$

Заметим, что, используя (3.226), можно формулу (3.231) переписать в виде

$$\frac{M_{2min}}{c^2} = \frac{2}{\sum\limits_{k=1}^{\infty}\left(1 + \frac{N_0}{c}\mu_k\right)^{-2}}. \qquad (3.236')$$

При $N_0 \to 0$ знаменатель в (3.236') неограниченно возрастает (μ_k от N_0 не зависит) и соответственно $M_{2min} \to 0$. Это указывает на возможность оценки параметра без смещения и с нулевой дисперсией при отсутствии аддитивного белого шума, что соответствует рассмотренному выше сингулярному случаю [см. (3.130)].

Рассмотрим теперь следующий специальный класс нелинейных оценок неизвестного параметра c корреляционной

функции нормального случайного процесса (рис. 21):

$$\hat{c} = \int_{-T}^{T} \int_{-T}^{T} h(t, u)\, x(t)\, x(u)\, dt\, du - N_0 \int_{-T}^{T} h(t, t)\, dt. \qquad (3.237)$$

Нетрудно убедиться, что эти оценки *несмещенные* при условии

$$\int_{-T}^{T} \int_{-T}^{T} h(t, u)\, R(t, u)\, dt\, du = 1, \qquad (3.238)$$

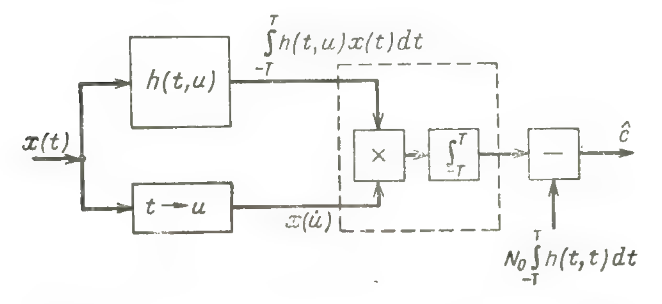


Рис. 21. Блок-схема устройства для оценки амплитудного множителя корреляционной функции.

так как с учетом (3.218)

$$m_1\{\hat{c}\} = \int_{-T}^{T} \int_{-T}^{T} h(t, u)\, B_\xi(t, u)\, dt\, du - N_0 \int_{-T}^{T} h(t, t)\, dt = $$

$$= c \int_{-T}^{T} \int_{-T}^{T} h(t, u)\, R(t, u)\, dt\, du = c.$$

Дисперсия оценки (3.237) равна

$$M_2\{\hat{c}\} = m_1\{\hat{c}^2\} - c^2 =$$

$$= \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} h(t, u)\, h(v, y)\, m_1\{x(t)\, x(u)\, x(v)\, x(y)\} \times$$

$$\times dt\, du\, dv\, dy - 2N_0 \int_{-T}^{T} h(t, t)\, dt \int_{-T}^{T} \int_{-T}^{T} h(t, u)\, B_\xi(t, u)\, dt\, du +$$

$$+ N_0^2 \left(\int_{-T}^{T} h(t, t)\, dt \right)^2 - c^2.$$

Используя известное выражение для смешанного момента четвертого порядка нормального случайного процесса (см. задачу 4.5 в первой книге), получаем с учетом (3.218) и нормировки (3.238)

$$M_2\{\hat{c}\} = 2N_0^2 \int_{-T}^{T}\int_{-T}^{T} h(t, v)\, h(v, t)\, dt\, dv +$$

$$+ 4N_0 c \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} h(t, u)\, h(t, y)\, R(u, y)\, du\, dy\, dt +$$

$$+ c^2 \int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T} h(t, u)\, h(v, y)\, [R(t, v)\, R(u, y) +$$

$$+ R(t, y)\, R(u, v)]\, dt\, du\, dv\, dy. \qquad (3.239)$$

Определение весовой функции $h(t, u)$, для которой дисперсия $M_2\{\hat{c}\}$ минимальна при условии (3.238), представляет сложную задачу вариационного исчисления. Однако нетрудно решить эту задачу для предельных случаев малых и больших отношений $\frac{N_0}{cT}$.

Если $\frac{N_0}{cT} \gg 1$, то сформулированная выше задача сводится к минимизации функционала [первого члена в (3.239)]

$$\int_{-T}^{T}\int_{-T}^{T} h(t, v)\, h(v, t)\, dt\, dv$$

при условии (3.238) Решение имеет простой вид:

$$h(t, u) = \frac{R(t, u)}{\int_{-T}^{T}\int_{-T}^{T} R^2(u, v)\, du\, dv}. \qquad (3.240)$$

Следовательно, в рассматриваемом случае

$$M_2\{\hat{c}\} \sim 2N_0^2 \frac{\int_{-T}^{T}\int_{-T}^{T} R^2(t, v)\, dt\, dv}{\left(\int_{-T}^{T}\int_{-T}^{T} R^2(u, v)\, du\, dv\right)^2} =$$

$$= 2N_0^2 \left[\int_{-T}^{T}\int_{-T}^{T} R^2(u, v)\, du\, dv\right]^{-1}, \qquad (3.241)$$

что по существу не отличается от (3.235).

Итак, при $\frac{N_0}{cT} \gg 1$ несмещенная оценка

$$\hat{c} = \frac{1}{\int\limits_{-T}^{T}\int\limits_{-T}^{T} R^2(u, v)\,du\,dv} \left[\int\limits_{-T}^{T}\int\limits_{-T}^{T} R(t, u)\,x(t)\,x(u)\,dt\,du - N_0 \int\limits_{-T}^{T} R(t, t)\,dt\right] \quad (3.242)$$

представляет *асимптотически эффективную* оценку.

Для стационарного нормального процесса (3.242) может быть записано в виде

$$\hat{c} = \frac{1}{\int\limits_{-2T}^{2T} \left(1 - \frac{|\tau|}{2T}\right) R^2(\tau)\,d\tau} \times \left[\frac{1}{2T}\int\limits_{-T}^{T}\int\limits_{-T}^{T} R(t-u)\,x(t)\,x(u)\,dt\,du - N_0\right]. \quad (3.243)$$

В другом предельном случае, когда $\frac{N_0}{cT} \to 0$,

$$M_2\{\hat{c}\} \to c^2 \int\limits_{-T}^{T}\int\limits_{-T}^{T}\int\limits_{-T}^{T}\int\limits_{-T}^{T} h(t, u)\,h(v, y) \times [R(t, v)\,R(u, y) + R(t, y)\,R(u, v)]\,dt\,du\,dv\,dy > 0. \quad (3.244)$$

Таким образом, в классе рассматриваемых несмещенных оценок дисперсия оценки даже при $\frac{N_0}{cT} \to 0$ (когда возможна сингулярность) ограничена снизу величиной, отличной от нуля.

3.6.8. Оценка корреляционной функции. В предшествующих разделах рассматривались оценки неизвестных параметров функции распределения случайного процесса. В некоторых ситуациях для заданного вида распределений неизвестными оказываются детерминированные функции, характеризующие распределение. К числу таких функций относятся, прежде всего, корреляционные функции и энер-

гетический спектр стационарного (в широком смысле) случайного процесса. В этом и следующем разделах рассматриваются оценки указанных функций по одной реализации $x(t)$ процесса, наблюдаемой на интервале $(-T, T)$.

Известно (см. § 4.1.6 в первой книге), что для эргодических случайных процессов временнóе среднее

$$B_T(\tau) = \frac{1}{2T} \int_{-T}^{T} x(t)\, x(t+\tau)\, dt \tag{3.245}$$

сходится при $T \to \infty$ по вероятности к корреляционной функции $B(t)$ случайного процесса. При дополнительных предположениях (например, нормальности) имеет место сходимость $B_T(\tau)$ к $B(\tau)$ в среднеквадратическом (см. задачу 4.2 в первой книге).

Естественно, в качестве оценки неизвестной корреляционной функции $B(\tau)$ случайного процесса можно взять выборочную временнýю корреляционную функцию (3.245). Однако она неудобна в том смысле, что для ее вычисления при любом $|\tau| \leqslant T$ необходимо иметь значения реализации $x(t)$ вне заданного интервала наблюдения $(-T, T)$. Поэтому в качестве оценки корреляционной функции $B(\tau)$ принимают функцию

$$\hat{B}(\tau) = B_T^*(\tau) =$$

$$= \frac{1}{2T - |\tau|} \int_{-T+\frac{|\tau|}{2}}^{T-\frac{|\tau|}{2}} x\left(t - \frac{\tau}{2}\right) x\left(t + \frac{\tau}{2}\right) dt, \quad |\tau| \leqslant T. \tag{3.246}$$

Оценка (3.246), так же как и (3.245), состоятельная и несмещенная. Последнее следует из того, что среднее по множеству реализаций

$$m_1\{B_T^*(\tau)\} =$$

$$= \frac{1}{2T - |\tau|} \int_{-T+\frac{|\tau|}{2}}^{T-\frac{|\tau|}{2}} m_1\left\{x\left(t - \frac{\tau}{2}\right) x\left(t + \frac{\tau}{2}\right)\right\} dt = B(\tau).$$

(3.247)

Для того чтобы определить дисперсию оценки (3.246), найдем сначала второй момент случайной величины $B_T^*(\tau)$:

$$m_2\{B_T^*(\tau)\} = \frac{1}{(2T-|\tau|)^2} \int\limits_{-T+\frac{|\tau|}{2}}^{T-\frac{|\tau|}{2}} \int\limits_{-T+\frac{|\tau|}{2}}^{T-\frac{|\tau|}{2}} m_1\left\{x\left(u-\frac{\tau}{2}\right)\times\right.$$

$$\left.\times x\left(u+\frac{\tau}{2}\right) x\left(v-\frac{\tau}{2}\right) x\left(v+\frac{\tau}{2}\right)\right\} du\, dv. \quad (3.248)$$

Из предположения о стационарности следует, что четвертый момент $m_1\left\{x\left(u-\frac{\tau}{2}\right) x\left(u+\frac{\tau}{2}\right) x\left(v-\frac{\tau}{2}\right) x\left(v+\frac{\tau}{2}\right)\right\}$ зависит только от двух аргументов: τ и $v-u$. Обозначая этот момент символом $M(\tau, v-u)$, перепишем (3.248) в виде *)

$$m_2\{B_T^*(\tau)\} = \frac{1}{(2T-|\tau|)^2} \int\limits_{-T+\frac{|\tau|}{2}}^{T-\frac{|\tau|}{2}} \int\limits_{-T+\frac{|\tau|}{2}}^{T-\frac{|\tau|}{2}} M(\tau, v-u)\, du\, dv =$$

$$= \frac{1}{(2T-|\tau|)^2} \int\limits_{-2T+|\tau|}^{2T-|\tau|} (2T-|\tau|-|z|)\, M(\tau, z)\, dz.$$

Так как

$$\frac{1}{(2T-|\tau|)^2} \int\limits_{-2T+|\tau|}^{2T-|\tau|} (2T-|\tau|-|z|)\, dz = 1,$$

то с учетом (3.247)

$$M_2\{B_T^*(\tau)\} = m_2\{B_T^*(\tau)\} - B^2(\tau) =$$

$$= \frac{1}{(2T-|\tau|)^2} \int\limits_{-2T+|\tau|}^{2T-|\tau|} (2T-|\tau|-|z|)\,[M(\tau, z) - B^2(\tau)]\, dz. \quad (3.249)$$

Для *нормального* стационарного случайного процесса момент четвертого порядка равен (см. задачу 4.5 в первой книге)

$$M(\tau, z) = B^2(\tau) + B^2(z) + B(\tau+z)\, B(\tau-z). \quad (3.250)$$

---

*) Переход от двойного интеграла к одномерному аналогичен преобразованию (4.60) в (4.61), приведенному в первой книге.

Подставляя (3.250) в (3.249), получим для дисперсии оценки корреляционной функции нормального процесса следующее выражение:

$$M_2\{B_T^*(\tau)\} = \frac{2}{2T-|\tau|} \int\limits_0^{2T-|\tau|} \left(1-\frac{z}{2T-|\tau|}\right) \times$$

$$\times [B^2(z) + B(\tau+z) B(\tau-z)]\, dz. \qquad (3.251)$$

Распределение вероятностей оценки (3.246) можно найти методами, указанными в § 9.3 первой книги. Из приведенных там результатов следует, в частности, что при $T \to \infty$ эта оценка асимптотически нормальна со средним $B(\tau)$ и дисперсией:

$$M_2\{B_T^*(\tau)\} \sim \frac{1}{T} \int\limits_0^{\infty} [B^2(z) + B(\tau+z) B(\tau-z)]\, dz. \qquad (3.252)$$

При $\tau = 0$ из (3.246) находим несмещенную оценку средней мощности стационарного случайного процесса

$$m_2^* = B_T^*(0) = \frac{1}{2T} \int\limits_{-T}^{T} x^2(t)\, dt, \qquad (3.253)$$

дисперсия которой для нормального процесса согласно (3.251) равна

$$M_2\{m_2^*\} = \frac{2}{T} \int\limits_0^{2T} \left(1-\frac{z}{2T}\right) B^2(z)\, dz. \qquad (3.254)$$

При $T \to \infty$ оценка (3.253) средней мощности асимптотически нормальная со средним $B(0)$ и дисперсией

$$M_2\{m_2^*\} \sim \frac{2}{T} \int\limits_0^{\infty} B^2(z)\, dz. \qquad (3.255)$$

Заметим, что оценка (3.253) является частным случаем (3.237) при $N_0 = 0$ и $h(t, u) = \frac{1}{2T}\delta(t-u)$. При этом формула (2.239) переходит в (2.254).

Приведенные выше результаты нетрудно распространить на оценку взаимной корреляционной функции $B_{\xi\eta}(\tau)$ двух

стационарных и стационарно связанных процессов $\xi(t)$ и $\eta(t)$. Аналогично (3.246) имеем

$$B^*_{\xi\eta T}(\tau) = \frac{1}{2T-|\tau|} \int\limits_{-T+\frac{|\tau|}{2}}^{T-\frac{|\tau|}{2}} x\left(t-\frac{\tau}{2}\right) y\left(t+\frac{\tau}{2}\right) dt, \quad |\tau| \leqslant T, \tag{3.256}$$

где $x(t)$ и $y(t)$ — реализации процессов $\xi(t)$ и $\eta(t)$ соответственно, наблюдаемые на интервале времени $(-T, T)$.

Оценка (3.256) несмещенная, а ее дисперсия равна [ср. (3.249)]

$$M_2\{B^*_{\xi\eta T}(\tau)\} = \\ = \frac{1}{2T-|\tau|} \int\limits_{-2T+|\tau|}^{2T-|\tau|} \left(1-\frac{|z|}{2T-|\tau|}\right) [M_{\xi\eta}(\tau, z) - B^2_{\xi\eta}(\tau)]\, dz, \tag{3.257}$$

где

$$M_{\xi\eta}(\tau, z) = m_1\left\{x\left(u-\frac{\tau}{2}\right) y\left(u+\frac{\tau}{2}\right) x\left(v-\frac{\tau}{2}\right) y\left(v+\frac{\tau}{2}\right)\right\}. \tag{3.258}$$

3.6.9. Оценка энергетического спектра. Примем сначала в качестве оценки непрерывного*) энергетического спектра стационарного случайного процесса по его реализации $x(t)$ на интервале $(-T, T)$ величину (см. § 4.7.3 в первой книге) **)

$$\hat{F}(\omega) = G_T(\omega) = \frac{1}{T} |Z_T(i\omega)|^2, \tag{3.259}$$

где

$$Z_T(i\omega) = \int\limits_{-T}^{T} x(t)\, e^{-i\omega t}\, dt. \tag{3.260}$$

---

*) Предполагается, таким образом, что среднее процесса равно нулю и он не содержит квазидетерминированных составляющих.

**) В статистической литературе оценку $G_T(\omega)$ называют *периодограммой*.

Нетрудно убедиться, что оценка (3.259) *смещенная.* Используя (4.61) первой книги, запишем

$$m_1\{G_T(\omega)\} = F_T(\omega) = 4\int_0^T \left(1-\frac{\tau}{T}\right) B(\tau)\cos\omega\tau\, d\tau =$$

$$= F(\omega) - 4\left(\frac{1}{T}\int_0^T \tau B(\tau)\cos\omega\tau\, d\tau + \int_T^\infty B(\tau)\cos\omega\tau\, d\tau\right).$$

Следовательно, смещение оценки энергетического спектра при заданном T равно

$$F(\omega) - F_T(\omega) =$$

$$= \frac{4}{T}\int_0^T \tau B(\tau)\cos\omega\tau\, d\tau + 4\int_T^\infty B(\tau)\cos\omega\tau\, d\tau, \quad (3.261)$$

где $B(\tau)$ — корреляционная функция процесса.

Однако из (3.261) (как, впрочем, и из определения энергетического спектра) следует, что оценка (3.259) при $T\to\infty$ *асимптотически несмещенная*, так как

$$\lim_{T\to\infty} F_T(\omega) = F(\omega),$$

при условии, что $\int_0^\infty \tau B(\tau)\, d\tau < \infty$.

Найдем дисперсию оценки (3.259) энергетического спектра *нормального* случайного процесса. Для этого определим сначала ее второй момент:

$$m_2\{G_T(\omega)\} = m_1\{G_T^2(\omega)\} =$$

$$= \frac{1}{T^2}\int_{-T}^T\int_{-T}^T\int_{-T}^T\int_{-T}^T m_1\{x(t)\,x(y)\,x(u)\,x(v)\}\times$$

$$\times e^{-i\omega(t-y+u-v)}\, dt\, dy\, du\, dv;$$

так как для нормального процесса (см. задачу 4.5 в первой книге)

$$m_1\{x(t)\,x(y)\,x(u)\,x(v)\} = B(y-t)\,B(v-u) +$$

$$+ B(u-t)\,B(v-y) + B(v-t)\,B(u-y), \quad (3.262)$$

то (см. (4.60) в первой книге)

$$m_2\{G_T(\omega)\} = \frac{1}{T^2}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}\int_{-T}^{T}[B(y-t)B(v-u)+$$

$$+B(u-t)B(v-y)+B(v-t)B(u-y)]\,e^{-i\omega(t-y+u-v)}\times$$

$$\times\,dt\,dy\,du\,dv = 2F_T^2(\omega)+\frac{1}{T^2}\int_{-T}^{T}\int_{-T}^{T}B(t-u)\times$$

$$\times\,e^{-i\omega(t+u)}\,dt\,du\int_{-T}^{T}\int_{-T}^{T}B(v-y)\,e^{i\omega(y+v)}\,dy\,dv$$

или

$$m_2\{G_T(\omega)\} = 2F_T^2(\omega)+\frac{1}{T^2}\left|\int_{-T}^{T}\int_{-T}^{T}B(u-t)\,e^{-i\omega(t+u)}\,dt\,du\right|^2. \tag{3.263}$$

Используя (3.263), находим искомую дисперсию оценки:

$$M_2\{G_T(\omega)\} = m_2\{G_T(\omega)\} - m_1^2\{G_T(\omega)\} =$$

$$= F_T^2(\omega)+\frac{1}{T^2}\left|\int_{-T}^{T}\int_{-T}^{T}B(u-t)\,e^{-i\omega(t+u)}\,dt\,du\right|^2. \tag{3.264}$$

Так как при $T\to\infty$ и $\omega\neq 0$

$$\frac{1}{T^2}\left|\int_{-T}^{T}\int_{-T}^{T}B(u-t)\,e^{-i\omega(t+u)}\,dt\,du\right|^2\sim$$

$$\sim F_T^2(\omega)\left(\frac{\sin 2\omega T}{2\omega T}\right)^2, \tag{3.265}$$

то второе слагаемое в (3.264) при $T\to\infty$ и $\omega\neq 0$ стремится к нулю и, следовательно,

$$\lim_{T\to\infty} M_2\{G_T(\omega)\} = F^2(\omega). \tag{3.266}$$

При $\omega=0$ и $T\to\infty$ из (3.264) непосредственно следует

$$\lim_{T\to\infty} M_2\{G_T(0)\} = F^2(0).$$

Таким образом, для всех ω, при которых $F(\omega)>0$, дисперсия оценки энергетического спектра *не стремится к нулю* при $T\to\infty$. Иначе говоря, оценка не сходится в среднеквадратическом к оцениваемому спектру при $T\to\infty$.

Сохранив предположение о нормальности процесса, можно найти распределение вероятностей оценки (3.259) его энергетического спектра. Эта оценка представляет сумму квадратов нормальных случайных величин

$$\frac{1}{\sqrt{T}}\int_{-T}^{T} x(t)\cos\omega t\,dt,\qquad \frac{1}{\sqrt{T}}\int_{-T}^{T} x(t)\sin\omega t\,dt,$$

средние значения которых равны нулю, а их дисперсии и ковариация равны соответственно

$$\sigma_{Tc}^2(\omega)=m_1\left\{\frac{1}{T}\left(\int_{-T}^{T} x(t)\cos\omega t\,dt\right)^2\right\}=$$

$$=\frac{1}{T}\int_{-T}^{T}\int_{-T}^{T} B(v-u)\cos\omega v\cos\omega u\,dv\,du,\qquad(3.267)$$

$$\sigma_{T_s}^2(\omega)=m_1\left\{\frac{1}{T}\left(\int_{-T}^{T} x(t)\sin\omega t\,dt\right)^2\right\}=$$

$$=\frac{1}{T}\int_{-T}^{T}\int_{-T}^{T} B(v-u)\sin\omega v\sin\omega u\,dv\,du,\qquad(3.268)$$

$$m_T(\omega)=m_1\left\{\frac{1}{T}\int_{-T}^{T} x(t)\cos\omega t\,dt\int_{-T}^{T} x(t)\sin\omega t\,dt\right\}=$$

$$=\frac{1}{T}\int_{-T}^{T}\int_{-T}^{T} B(v-u)\cos\omega v\sin\omega v\,dv\,du.\qquad(3.269)$$

Используя формулу (3.116) первой книги, запишем теперь выражение характеристической функции оценки энергетического спектра:

$$\Theta_{G_T}(v)=\frac{1}{2\pi\sigma_{Tc}\sigma_{Ts}\sqrt{1-R_T^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left\{-\frac{1}{2(1-R_T^2)}\times\right.$$

$$\left.\times\left(\frac{x_1^2}{\sigma_{Tc}^2}-2R_T\frac{x_1x_2}{\sigma_{Tc}\sigma_{Ts}}+\frac{x_2^2}{\sigma_{Ts}^2}\right)+iv(x_1^2+x_2^2)\right\}dx_1dx_2=$$

$$=\frac{\sqrt{1-R_T^2}}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left\{\frac{1}{2}(y_1^2-2R_Ty_1y_2+y_2^2)+\right.$$

$$\left.+iv(1-R_T^2)(\sigma_{Tc}^2y_1^2+\sigma_{Ts}^2y_2^2)\right\}dy_1\,dy_2,$$

где обозначено

$$R_T = \frac{m_T}{\sigma_{Tc}\sigma_{Ts}}. \tag{3.270}$$

Вычисление интеграла производим согласно приложению II первой книги. Получим

$$\Theta_{G_T}(v) = \sqrt{1 - R_T^2}\,[1 - 2iv\,(1 - R_T^2)(\sigma_{Tc}^2 + \sigma_{Ts}^2) - $$
$$- 4v^2(1 - R_T^2)^2\,\sigma_{Tc}^2\sigma_{Ts}^2 - R_T^2]^{-\frac{1}{2}}$$

или

$$\Theta_{G_T}(v) = [1 - 2iv\,(\sigma_{Tc}^2 + \sigma_{Ts}^2) - 4v^2(1 - R_T^2)\,\sigma_{Tc}^2\sigma_{Ts}^2]^{-\frac{1}{2}}. \tag{3.271}$$

Обратным преобразованием Фурье *) находим искомое распределение оценки энергетического спектра нормального случайного процесса:

$$W_{G_T}(z) = \frac{1}{2\sigma_{Tc}\sigma_{Ts}\sqrt{1 - R_T^2}} \exp\left[-\frac{\sigma_{Tc}^2 + \sigma_{Ts}^2}{4\sigma_{Tc}^2\sigma_{Ts}^2(1 - R_T^2)}\,z\right] \times$$
$$\times I_0\left[\frac{\sqrt{(\sigma_{Tc}^2 - \sigma_{Ts}^2)^2 + 4\sigma_{Tc}^2\sigma_{Ts}^2R_T^2}}{4\sigma_{Tc}^2\sigma_{Ts}^2(1 - R_T^2)}\,z\right], \quad z \geqslant 0. \tag{3.272}$$

Согласно определению энергетического спектра

$$\lim_{T\to\infty} \sigma_{Tc}^2(\omega) = \lim_{T\to\infty} \sigma_{Ts}^2(\omega) = \frac{1}{2}F(\omega), \tag{3.273}$$

$$\lim_{T\to\infty} m_T(\omega) = 0. \tag{3.274}$$

Из (3.272) следует, что при $T \to \infty$ оценка энергетического спектра имеет в пределе *экспоненциальное распределение* вероятностей

$$W_{G_T}(z) \sim \frac{1}{F(\omega)}\,e^{-\frac{z}{F(\omega)}}, \quad z > 0 \tag{3.275}$$

с параметром $F(\omega)$ [т. е. среднее равно $F(\omega)$, а дисперсия $F^2(\omega)$ находится в полном согласии с (3.266)].

Хотя рассмотренная выше оценка $\hat{F}(\omega) = G_T(\omega)$ энергетического спектра случайного процесса асимптотически несмещенная, она обладает тем недостатком, что увеличение

*) См., например, [5, т. 2, стр. 328].

времени наблюдения над реализацией случайного процесса не позволяет достигнуть желаемой точности, так как даже при $T \to \infty$ дисперсия оценки остается конечной. Поэтому рассмотрим также класс *смещенных* оценок с дисперсией, убывающей при увеличении времени наблюдения.

Выберем в качестве оценки непрерывного энергетического спектра стационарного случайного процесса следующую случайную величину, представляющую функционал от наблюдаемой на интервале $(-T, T)$ реализации процесса:

$$\hat{F}(\omega) = g_T(\omega) = \\ = \frac{1}{T} \int_{-T}^{T} \left| \int_{-T}^{T} x(u)\, h(t-u)\, e^{-i\omega u}\, du \right|^2 dt, \qquad (3.276)$$

где $h(\tau)$ — весовая функция, которую можно трактовать как импульсную переходную функцию некоторого линейного фильтра. Оценка (3.276) получается усреднением по времени на интервале наблюдения суммы квадратов процессов на выходе упомянутого фильтра, когда на вход действуют процессы $x(t) \cos \omega t$ и $x(t) \sin \omega t$.

Вводя преобразования Фурье $Z_T(i\omega)$ для реализации $x(t)$ и $k(i\omega)$ для импульсной переходной функции $h(\tau)$ (см. (5.5′) в первой книге), перепишем (3.276) в виде

$$g_T(\omega) = \frac{1}{(2\pi)^4} \int_{-T}^{T} \int_{-T}^{T} \int_{-T}^{T} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{T} Z_T(i\omega_1) \times \\ \times Z_T(i\omega_2)\, k(i\omega_3)\, k(i\omega_4) \exp[iu(\omega_1 - \omega_3 - \omega) + \\ + iv(\omega_2 - \omega_4 + \omega) + it(\omega_3 + \omega_4)]\, du\, dv\, dt\, d\omega_1\, d\omega_2\, d\omega_3\, d\omega_4. \qquad (3.277)$$

Усредняя обе части (3.277) по множеству реализаций, переходя к пределу при $T \to \infty$ и используя фильтрующее свойство дельта-функции (см. приложение III в первой книге)

$$\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega u}\, du,$$

получаем

$$m_1\{g_T(\omega)\} \sim \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega + \nu)\, C^2(\nu)\, d\nu, \qquad (3.278)$$

где

$$C^2(\omega) = |k(i\omega)|^2.$$

Из (3.278) видно, что оценка (3.276) смещенная. С учетом нормировки характеристики фильтра

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} C^2(\nu)\,d\nu = \int_{-\infty}^{\infty} h^2(\tau)\,d\tau = 1 \tag{3.279}$$

асимптотическое смещение оценки можно представить в виде

$$m_1\{g_T(\omega)\} - F(\omega) \sim \frac{1}{2\pi}\int_{-\infty}^{\infty} [F(\nu) - F(\omega)]\, C^2(\nu-\omega)\,d\nu. \tag{3.280}$$

При дополнительном предположении нормального распределения процесса $x(t)$ можно найти асимптотическое выражение (при $T \to \infty$) для дисперсии оценки (3.276) [15]:

$$M_2\{g_T(\omega)\} \sim \frac{1}{T} F^2(\omega)\int_{-\infty}^{\infty} C^4(\omega)\,d\omega. \tag{3.281}$$

При сглаживающем фильтре типа RC-интегратора (см. § 5.2.3 в первой книге) с учетом нормировки (3.279)

$$C^2(\omega) = \frac{2/\alpha}{1+\left(\frac{\omega}{\alpha}\right)^2}, \qquad \alpha = \frac{1}{RC},$$

из (3.281) получаем

$$M_2\{g_T(\omega)\} \sim \frac{2\pi}{\alpha T} F^2(\omega), \tag{3.282}$$

т. е. для получения достаточно малого значения дисперсии оценки время наблюдения T должно быть много больше постоянной времени $\frac{1}{\alpha} = RC$ фильтра (иначе говоря, произведение ширины полосы фильтра и времени наблюдения должно быть много больше единицы).

Задачи

3.1. Пусть $s(t)$ — детерминированная функция, спектр (преобразование Фурье) которой непрерывен и ограничен полосой частот (ω_1, ω_2). Показать, что для этого случая справедлива следующая интерполяционная формула:

$$s(t) = \sum_{n=-\infty}^{\infty} A\left(\frac{\pi n}{\Delta}\right) \frac{\sin\left[\Delta\left(t - \frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t - \frac{\pi n}{\Delta}\right)} \times$$

$$\times \cos\left[\omega_0\left(t - \frac{\pi n}{\Delta}\right) - \varphi\left(\frac{\pi n}{\Delta}\right)\right], \tag{1}$$

где

$$\Delta = \frac{\omega_2 - \omega_1}{2}; \quad \omega_0 = \frac{\omega_2 + \omega_1}{2}; \tag{2}$$

$$A^2(t) = s^2(t) + \sigma^2(t); \tag{3}$$

$$\varphi(t) = \operatorname{arctg} \frac{\sigma(t)}{s(t)}; \tag{3'}$$

$\sigma(t)$ — преобразование Гильберта от $s(t)$ (см. приложение VII в первой книге).

3.2. Показать, что площадь под детерминированной функцией $f(t)$, имеющей непрерывный спектр, тождественно равный нулю при $|\omega| > \Delta$, равна

$$\int_{-\infty}^{\infty} f(t)\,dt = \frac{\pi}{\Delta} \sum_{n=-\infty}^{\infty} f\left(\frac{\pi n}{\Delta}\right). \tag{4}$$

Используя (4) и (4.67') в первой книге, показать, что время корреляции стационарного в широком смысле случайного процесса, энергетический спектр которого непрерывен и ограничен полосой частот $(-\Delta, \Delta)$, равно

$$\tau_0 = \frac{\pi}{4\Delta} \sum_{n=-\infty}^{\infty} R\left(\frac{\pi n}{\Delta}\right), \tag{5}$$

где $R(\tau)$ — коэффициент корреляции процесса.

3.3. Имея в виду, что амплитудный спектр функции $\frac{\sin \Delta (t - t_0)}{\Delta (t - t_0)}$ при любом фиксированном t_0 обращается в нуль при $|\omega| > \Delta$ [см. (3.11)], доказать справедливость равенства

$$\frac{\sin \Delta (t - t_0)}{\Delta (t - t_0)} = \sum_{n=-\infty}^{\infty} \frac{\sin \Delta \left(t_0 - \frac{\pi n}{\Delta}\right)}{\Delta \left(t_0 - \frac{\pi n}{\Delta}\right)} \frac{\sin \Delta \left(t - \frac{\pi n}{\Delta}\right)}{\Delta \left(t - \frac{\pi n}{\Delta}\right)}. \tag{6}$$

Доказать (6), используя ортогональное разложение корреляционной функции (см. § 3.3.2).

3.4. Для любого фиксированного t_0 получить следующее обобщение формулы (3.4):

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(t_0 + \frac{\pi n}{\Delta}\right) \frac{\sin\left[\Delta\left(t - t_0 - \frac{\pi n}{\Delta}\right)\right]}{\Delta\left(t - t_0 - \frac{\pi n}{\Delta}\right)}. \tag{7}$$

3.5. Показать, что оценка максимального правдоподобия амплитуды a квазидетерминированного процесса $as(t)$ по наблюдаемой на интервале $(-T, T)$ реализации $x(t)$ суммы этого процесса и стационарного нормального марковского процесса с дисперсией σ^2

и коэффициентом корреляции $R(\tau) = \exp(-\mu|\tau|)$ имеет следующий вид:

$$\hat{a} = \left\{\left[s(-T) - \frac{1}{\mu}s'(-T)\right]x(-T) + \left[s(T) - \frac{1}{\mu}s'(T)\right]x(T) + \right.$$

$$\left. + \mu\int_{-T}^{T}\left[s(t) - \frac{1}{\mu}s''(t)\right]x(t)\,dt\right\}\left[s^2(-T) + s^2(T) + \right.$$

$$\left. + \mu\int_{-T}^{T}\left\{s^2(t) + \frac{1}{\mu^2}[s'(t)]^2\right\}dt\right]^{-1}. \quad (8)$$

Доказать, что оценка (8) несмещенная и что ее дисперсия равна [ср. (3.112)]

$$M_2\{\hat{a}\} = 2\sigma^2\left[s^2(-T) + s^2(T) + \right.$$

$$\left. + \mu\int_{-T}^{T}\left\{s^2(t) + \frac{1}{\mu^2}[s'(t)]^2\right\}dt\right]^{-1}. \quad (9)$$

Рассмотреть частный случай $s(t) \equiv 1$ и убедиться, что при этом

$$\hat{a} = \frac{x(-T) + x(T) + \mu\int_{-T}^{T}x(t)\,dt}{2(1 + \mu T)}, \quad (10)$$

$$M_2\{\hat{a}\} = \frac{\sigma^2}{1 + \mu T} \quad (11)$$

и при $\mu T \gg 1$

$$\hat{a} \sim \frac{1}{2T}\int_{-T}^{T}x(t)\,dt, \quad (12)$$

$$M_2\{\hat{a}\} \sim \frac{\sigma^2}{\mu T}. \quad (13)$$

Объяснить причину совпадения (12) с (3.169).

3.6. Рассмотреть задачу § 3.6.5 в предположении, что результаты наблюдения представлены не реализацией $x(t) = s(t;\ \vartheta_1,\ \ldots,\ \vartheta_m) + \xi(t)$, а дискретной выборкой $\mathbf{x} = (x_1,\ \ldots\ x_N)$, где $x_i = x(t_i)$ и t_i — моменты времени, принадлежащие интервалу наблюдения $(-T,\ T)$. Показать, что оценки максимального правдоподобия параметров ϑ_k могут быть представлены в векторной форме [ср. (3.158)]

$$\hat{\boldsymbol{\vartheta}} = \mathbf{s}^{-1}\mathbf{X}, \quad (14)$$

где

$$\mathbf{X}=\mathbf{s}'\mathbf{k}^{-1}\mathbf{x}, \tag{15}$$

$$\mathbf{S}=\mathbf{s}'\mathbf{k}^{-1}\mathbf{s}. \tag{15'}$$

Здесь $\mathbf{k}$ — нормированная корреляционная матрица случайного слагаемого $\xi(t)$,

$$\mathbf{k}=\frac{N\mathbf{K}}{\mathrm{tr}\,\mathbf{K}}, \tag{16}$$

где $\mathrm{tr}\,\mathbf{K}$ — след корреляционной матрицы $\mathbf{K}$, а $\mathbf{s}$ — матрица размером $N\times m$ с линейно-независимыми (в алгебраическом смысле) столбцами, причем j-й столбец матрицы представляет вектор s_j с компонентами $s_j(t_1), \ldots, s_j(t_N)$; $j=1, \ldots, m$. Штрихом обозначена транспонированная матрица. Доказать, что оценка (14) несмещенная, а ее корреляционная матрица равна

$$\mathbf{M}=\frac{\mathrm{tr}\,\mathbf{K}}{N}\mathbf{S}^{-1}. \tag{17}$$

3.7. Показать, что в условиях задачи 3.6 совместная байесовская оценка совокупности параметров $\vartheta=(\vartheta_1, \ldots, \vartheta_m)$ при квадратичной функции потерь имеет следующий вид:

$$\hat{\boldsymbol{\vartheta}}=\frac{\int\limits_{\Omega}\boldsymbol{\vartheta}w_m(\boldsymbol{\vartheta})\,l_N(\mathbf{x}\,|\,\boldsymbol{\vartheta})\,d\boldsymbol{\vartheta}}{\int\limits_{\Omega}w_m(\boldsymbol{\vartheta})\,l_N(\mathbf{x}\,|\,\boldsymbol{\vartheta})\,d\boldsymbol{\vartheta}}, \tag{18}$$

где $w_m(\boldsymbol{\vartheta})$ — совместная априорная плотность вероятности оцениваемых параметров и Ω — область m-мерного пространства, на которой задана эта плотность, а функция $l_N(x\,|\,\boldsymbol{\vartheta})$ определяется из соотношения

$$l_N(\mathbf{x}\,|\,\boldsymbol{\vartheta})=\exp\left\{-\frac{N}{2\,\mathrm{tr}\,\mathbf{K}}(\mathbf{x}-\mathbf{s}\boldsymbol{\vartheta})'\,\mathbf{k}^{-1}(\mathbf{x}-\mathbf{s}\boldsymbol{\vartheta})\right\}. \tag{19}$$

3.8. Найти нижнюю границу M_{2min} дисперсии несмещенных оценок параметра σ^2 нормального стационарного марковского процесса, корреляционная функция которого представлена формулой (3.109), в аддитивном нормальном белом шуме со спектральной интенсивностью N_0 по реализации суммарного процесса на интервале $(-T, T)$. Доказать, что при $\frac{N_0}{\sigma^2 T}\gg 1$ имеет место асимптотическая формула

$$\frac{M_{2min}}{\sigma^4}\sim\frac{4\mu^2 T}{4\mu T+e^{-4\mu T}-1}\left(\frac{N_0}{\sigma^2 T}\right)^2. \tag{20}$$

Если время наблюдения много меньше времени корреляции марковского процесса ($\mu T\ll 1$), то показать, что из (20) следует

$$\frac{M_{2min}}{\sigma^4}\sim\frac{1}{2}\left(\frac{N_0}{\sigma^2 T}\right)^2. \tag{20'}$$

(См. также задачу 7.6 в первой книге.)

ЛИТЕРАТУРА

МОНОГРАФИИ

1. Вайнштейн Л. А., Зубаков В. Д. Выделение сигналов на фоне случайных помех. Изд-во «Советское радио», 1959, приложение III.
2. Гренандер У. Случайные процессы и статистические выводы. Пер. с англ., под ред. А. М. Яглома. Изд-во иностранной литературы, 1961.
3. Давенпорт В. Б., Рут В. Л. Введение в теорию случайных сигналов и шумов. Пер. с англ., под ред. Р. Л. Добрушина. Изд-во иностранной литературы, 1960, гл. 14 и приложение 2.
4. Дуб Дж. Л. Вероятностные процессы. Пер. с англ., под ред. А. М. Яглома. Изд-во иностранной литературы, 1956, гл. VII, § 9.
5. Миддлтон Д. Введение в статистическую теорию связи. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», т. I, 1961, гл. 4; т. II, 1962, гл. 16.
6. Серебрянников М. Г., Первозванский А. Л. Выявление скрытых периодичностей. Изд-во «Наука», 1965, гл. 5.
7. Хэннан Э. Анализ временны́х рядов. Пер. с англ., под ред. Ю. А. Розанова. Изд-во «Наука», 1964.
8. Харкевич А. А. Спектры и анализ. Гостехиздат, 1949

СТАТЬИ

9. Беляев Ю. К. Аналитические случайные процессы. «Теория вероятностей и ее применения», 1959, вып. 4.
10. Гаек Я. Линейная оценка среднего стационарного случайного процесса с выпуклой корреляционной функцией. Чехословацкий математический журнал, 1956, вып. 6.
11. Левин Б. Р., Шинаков Ю. С. Предельная форма байесовской оценки коэффициентов регрессии при нормальном нестационарном шуме. «Проблемы передачи информации», 1967, № 1.
12. Писаренко Б. Ф. К задаче обнаружения случайного сигнала на фоне шума. «Радиотехника и электроника», 1961, № 4.
13. Писаренко Б. Ф. О вычислении отношения правдоподобия для гауссовских процессов с рациональным спектром. «Теория вероятностей и ее приложения», 1965, вып. 2.
14. Blackman R. B., Tukey I. W. The measurement of Power Spectra trom point of fien of Communication Engineering. BSTJ, 1958, № 1, № 2.
15. Grenander U., Rosenblatt M. Some problems in estimating the spectrum of time series. Proc. Third Berkley Symp. Math. Stat., Los Angelos, 1956, v. 2.
16. Kelly E. J., Reed I. S., Root W. L. The detection of radar echocs in noise. I. J. Soc. Ind. Appl. Math., 1960, v. 8, № 2.
17. Hofstetter E. M. Some results on the stochastic signal parameter estimation problem. IEEE Trans, 1965, IT-11, № 3.

18. Levin M. J. Power spectrum parameter estimation. IEEE Trans, 1965, IT-11, № 1.
19. Reich E., Swerling P. Detection of sine wave in gaussian noise. J. Appl. Phys. Rev., 1953, v. 24, № 3.
20. Slepian D. Some comments on the detection ot gaussian signals in gaussian noise. IRE Trans., 1958, IT-4, № 2.
21. Kailath T. Some results on singular detection. Information and Control, 1966, v. 9, april.

ГЛАВА ЧЕТВЕРТАЯ

ФИЛЬТРАЦИЯ СЛУЧАЙНЫХ ПРОЦЕССОВ

4.1. ВВОДНЫЕ ЗАМЕЧАНИЯ

В предыдущей главе рассматривались оценки неизвестных параметров и функций, характеризующих распределение вероятности случайного процесса, по реализации этого процесса, наблюдаемой на конечном интервале времени. Другой практически важной задачей является отыскание преобразования наблюдаемой реализации для получения наилучшей в некотором смысле оценки значения случайного процесса в некоторый момент времени внутри или вне интервала наблюдения. Это преобразование в общем случае называют *фильтрацией* случайного процесса.

Два параграфа этой главы будут посвящены линейным преобразованиям. Вначале, пренебрегая физической осуществимостью линейного фильтра, используется неограниченный интервал наблюдения, а затем вводится ограничение этого интервала. Критерием качества оценки значения случайного процесса служит в одних случаях среднее по множеству реализаций квадрата ошибки, т. е. средний квадрат отклонения оценки от оцениваемого значения, а в других — отношение мощности полезного сигнала к мощности помехи. Рассматривается задача оценки *) значения случайного процесса $\xi(t)$ по наблюдаемой реализации $x(t)$ суммы этого процесса с другим процессом $\eta(t)$, коррелированным с $\xi(t)$, а на основании общих результатов исследуются некоторые частные случаи, представляющие интерес для радиотехнических приложений. Наряду с оценкой самого процесса $\xi(t)$ представляют также интерес оценки по указанной выше реализации $x(t)$, наблюдаемой на некотором конечном (или бесконечном) интервале времени,

---

*) Часто говорят о *выделении* одного процесса (сигнала) $\xi(t)$ из аддитивной смеси этого процесса с другим (помехой) $\eta(t)$.

линейных преобразований этого процесса. Рассматриваются следующие виды линейных преобразований: смещение момента оценки по оси времени, однократное и многократное дифференцирование и интегрирование, а также комбинации этих преобразований.

Вопросы нелинейной фильтрации по критерию минимума среднего квадрата ошибки рассматриваются в § 4.4. В данном параграфе рассматривается *один* из возможных подходов к решению задачи, связанный с характеризацией нелинейных систем при помощи бесконечной суммы интегралов Вольтерра. Получаемое при этом решение в точном виде не реализуемо, и поэтому исследуются приближения, которые, однако, в некоторых случаях могут существенно уменьшить среднеквадратическую ошибку, получаемую при использовании оптимальной линейной системы.

Другой подход к точному решению задачи оптимальной нелинейной фильтрации состоит в ограничении класса исследуемых процессов одномерными марковскими или компонентами многомерных марковских процессов [13, 21]. Наконец, иногда используются квазилинейные методы, основанные на использовании малого параметра (например, соответствующим образом определенного отношения сигнал/шум) [9, 10]. Последние два подхода в этой книге не рассмотрены, и для ознакомления с ними следует использовать указанные выше библиографические ссылки.

4.2. ЛИНЕЙНАЯ ФИЛЬТРАЦИЯ ПО КРИТЕРИЮ МИНИМУМА СРЕДНЕГО КВАДРАТА ОШИБКИ

4.2.1. Импульсная переходная функция оптимальной линейной системы. Пусть $x(t)$ — реализация суммарного процесса $\xi(t) + \eta(t)$, определенная для всех действительных значений t. В качестве оценки $\xi(t)$ примем профильтрованное значение реализации (см. § 5.2 в первой книге)

$$\hat{\xi}(t) = \int_{-\infty}^{\infty} h(t, \tau)\, x(\tau)\, d\tau, \qquad (4.1)$$

где $h(t, \tau)$ — импульсная переходная функция линейного фильтра (без ограничений, связанных с физической осуществимостью). Интегралы случайных процессов здесь

и далее предполагаются сходящимися в среднеквадратическом (см. первую книгу, § 3.5).

Заметим прежде всего, что оценка (4.1), вообще говоря, *смещенная*, так как

$$m_1\{\hat{\xi}(t)\} = \int_{-\infty}^{\infty} h(t, \tau)\,[a_\xi(\tau) + a_\eta(\tau)]\,d\tau \neq a_\xi(t),$$

где $a_\xi(t)$ и $a_\eta(t)$ — средние значения процессов $\xi(t)$ и $\eta(t)$.

Среднее значение (по множеству реализаций случайного процесса) квадрата ошибки $\xi(t) - \hat{\xi}(t)$ при использовании такой оценки равно

$$\varepsilon^2(t) = m_1\{[\xi(t) - \hat{\xi}(t)]^2\}. \tag{4.2}$$

Определим импульсную переходную функцию линейного фильтра, для которого средняя квадратическая ошибка ε^2 имеет величину, меньшую, чем для любой другой линейной системы *).

Подставим (4.1) в (4.2) и произведем элементарные преобразования. Тогда

$$\varepsilon^2(t) = m_1\Big\{\xi^2(t) - 2\int_{-\infty}^{\infty} h(t, \tau)\,x(\tau)\,\xi(t)\,d\tau +$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t, u)\,h(t, v)\,x(u)\,x(v)\,du\,dv\Big\} =$$

$$= m_1\{\xi^2(t)\} - 2\int_{-\infty}^{\infty} h(t, u)\,m_1\{x(\tau)\,\xi(t)\}\,d\tau +$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t, u)\,h(t, v)\,m_1\{x(u)\,x(v)\}\,du\,dv. \tag{4.3}$$

*) Подчеркнем еще раз, что оптимум ищется в классе линейных систем. Конечно, в некоторых случаях, применяя нелинейные системы, можно иметь меньшую ошибку, чем та, которую дает оптимальная линейная система. Если процессы $\xi(t)$ и $\eta(t)$ являются нормальными случайными процессами, то абсолютный оптимум, как будет показано ниже, всегда реализуется при помощи линейной системы. Рассматриваемая в этом разделе задача линейной фильтрации является обобщением на случайные процессы линейной средней квадратической регрессии (см. приложение V).

Пусть известны корреляционные функции $B_\xi(t_1, t_2)$, $B_\eta(t_1, t_2)$ процессов $\xi(t)$ и $\eta(t)$ и их взаимная корреляционная функция $B_{\xi\eta}(t_1, t_2)$. Так как

$$m_1\{x(u)\,x(v)\} = B_x(u, v) = B_\xi(u, v) + B_{\xi\eta}(u, v) + B_{\eta\xi}(u, v) + B_\eta(u, v), \tag{4.4}$$

$$m_1\{x(\tau)\,\xi(t)\} = B_{x\xi}(\tau, t) = B_\xi(\tau, t) + B_{\xi\eta}(\tau, t), \tag{4.5}$$

то из (4.3) — (4.5) находим

$$\varepsilon^2(t) = B_\xi(t, t) - 2\int_{-\infty}^{\infty} h(t, \tau)\,B_{x\xi}(\tau, t)\,d\tau + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t, u)\,h(t, v)\,B_x(u, v)\,du\,dv. \tag{4.6}$$

Из (4.6) следует, что среднеквадратическая ошибка линейной оценки $\hat{\xi}(t)$ зависит только от корреляционной и взаимной корреляционной функций процессов $\xi(t)$ и $\eta(t)$ и *не зависит* от более тонкой структуры этих процессов.

Покажем, что при заданных B_ξ, B_η, $B_{\xi\eta}$ наилучшую (в смысле принятого критерия) линейную фильтрацию процесса $\xi(t)$ из его аддитивной смеси с $\eta(t)$ будет осуществлять линейная система, импульсная переходная функция которой удовлетворяет интегральному уравнению *)

$$B_{x\xi}(\tau, t) = \int_{-\infty}^{\infty} h^*(t, y)\,B_x(\tau, y)\,dy. \tag{4.7}$$

Подставим (4.7) в (4.6) и покажем, что ошибка $\varepsilon^2(t)$ минимальна при условии $h(t, \tau) \equiv h^*(t, \tau)$. В самом деле, в результате указанной подстановки получим

$$\varepsilon^2(t) = B_\xi(t, t) - 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t, \tau)\,h^*(t, y)\,B_x(\tau, y)\,d\tau\,dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(t, u)\,h(t, v)\,B_x(u, v)\,du\,dv =$$

*) Здесь доказывается только достаточность, доказательство необходимости см., например, в [4].

$$= B_\xi(t, t) - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h^*(t, \tau)\, h^*(t, y)\, B_x(\tau, y)\, d\tau\, dy +$$

$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} B_x(u, v)\, [h(t, u) - h^*(t, u)] \times$$

$$\times [h(t, v) - h^*(t, v)]\, du\, dv. \qquad (4.8)$$

Так как только последний член в (4.8) содержит неизвестную функцию $h(t, u)$ и так как он неотрицательный (в силу положительной полуопределенности корреляционной функции $B_x(u, v)$; см. § 3.3.1), то минимальное значение ε^2 будет соответствовать такой линейной системе, передаточная функция которой обращает его в нуль. Как нетрудно видеть, это будет иметь место при условии $h(t, u) \equiv h^*(t, u)$, что и требовалось доказать.

Из (4.8) следует, что минимальное значение среднеквадратичной ошибки при использовании оптимальной линейной системы равно

$$\varepsilon^2_{min}(t) = B_\xi(t, t) - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h^*(t, u)\, h^*(t, v)\, B_x(u, v)\, du\, dv \qquad (4.9)$$

или с учетом (4.7)

$$\varepsilon^2_{min}(t) = B_\xi(t, t) - \int_{-\infty}^{\infty} h^*(t, u)\, B_{x\xi}(u, t)\, du. \qquad (4.9')$$

Используя (5.14) из первой книги, можно (4.9) переписать следующим образом:

$$\varepsilon^2_{min}(t) = B_\xi(t, t) - B_{\hat{\xi}}(t, t), \qquad (4.9'')$$

т. е. минимальный средний квадрат ошибки равен разности средних квадратов значений оцениваемого процесса и оценки *).

Величину минимальной среднеквадратической ошибки $\varepsilon^2_{min}(t)$ можно выразить также через интеграл от разности мгновенных энергетических спектров процесса $\xi(t)$ и его линейной оценки $\hat{\xi}(t)$ (см. стр. 219—220 в первой книге).

*) Из (4.9″) следует, что $B_{\hat{\xi}}(t, t) \leqslant B_\xi(t, t)$.

Из (4.9″) следует:

$$\varepsilon_{min}^2(t)=\frac{1}{2\pi}\int_0^\infty [\Phi_\xi(t,\ \omega)-\Phi_{\hat{\xi}}(t,\ \omega)]\,d\omega,$$

где $\Phi(t,\ \omega)$ — мгновенный энергетический спектр, а индекс при Φ указывает, какому процессу соответствует спектр.

Заметим, что для некогерентных *) процессов $\xi(t)$ и $\eta(t)$ уравнения

$$B_x(u,\ v)=B_\xi(u,\ v)+B_\eta(u,\ v), \qquad (4.10)$$

$$B_{x\xi}(u,\ v)=B_\xi(u,\ v) \qquad (4.10')$$

и основное интегральное уравнение (4.7) преобразуются к виду

$$B_\xi(\tau,\ t)=\int_{-\infty}^{\infty} h^*(t,\ y)\,[B_\xi(\tau,\ y)+B_\eta(\tau,\ y)]\,dy, \qquad (4.11)$$

а формула (4.9) может быть записана следующим образом:

$$\varepsilon_{min}^2(t)=B_\xi(t,\ t)-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h^*(t,\ u)\,h^*(t,\ v)\times$$
$$\times[B_\xi(u,\ v)+B_\eta(u,\ v)]\,du\,dv=B_\xi(t,\ t)-$$
$$-\int_{-\infty}^{\infty} h^*(t,\ u)\,B_\xi(u,\ t)\,du. \qquad (4.11')$$

Если процессы $\xi(t)$ и $\eta(t)$ стационарны и стационарно связаны (по крайней мере, в широком смысле), а фильтр представляет линейную систему с постоянными во времени параметрами, то интегральное уравнение (4.7) для определения импульсной переходной характеристики преобразуются к виду **)

$$B_{x\xi}(\tau)=\int_{-\infty}^{\infty} h^*(z)\,B_x(\tau-z)\,dz, \qquad (4.12)$$

*) Здесь используется определение некогерентности двух случайных процессов, данное на стр. 184 первой книги. Его, конечно, не следует отождествлять с понятиями «когерентный сигнал» и «некогерентный сигнал», широко используемыми в импульсной технике для различения ситуаций, когда фаза высокочастотного заполнения импульсов известна и когда эта фаза случайна.

**) Интегральное уравнение (4.12) часто называют уравнением Винера — Хопфа.

а для некогерентных процессов $\xi(t)$ и $\eta(t)$

$$B_{\xi}(\tau)=\int\limits_{-\infty}^{\infty} h^{*}(z)\,[B_{\xi}(\tau-z)+B_{\eta}(\tau-z)]\,dz. \qquad (4.12')$$

Величина среднего квадрата ошибки не зависит от момента времени, для которого выполняется оценка значения процесса $\xi(t)$, а ее минимальная величина в соответствии с (4.9) равна

$$\varepsilon^2_{min}=B_{\xi}(0)-\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} h^{*}(u)\,h^{*}(v)\,B_x(v-u)\,du\,dv=$$

$$=B_{\xi}(0)-\int\limits_{-\infty}^{\infty} h^{*}(v)\,B_{x\xi}(v)\,dv \qquad (4.13)$$

или

$$\varepsilon^2_{min}=B_{\xi}(0)-B_{\hat{\xi}}(0), \qquad (4.13')$$

т. е. минимальный средний квадрат ошибки равен разности средних мощностей оцениваемого процесса и оценки.

Выражая средние мощности через энергетические спектры процессов (см. (5.22) в первой книге), запишем величину ε^2_{min} в виде

$$\varepsilon^2_{min}=\frac{1}{2\pi}\int\limits_{0}^{\infty}[F_{\xi}(\omega)-F_x(\omega)\,|\,k^{*}(i\omega)\,|^2]\,d\omega, \qquad (4.14)$$

$$F_x(\omega)=F_{\xi}(\omega)+F_{\eta}(\omega)+F_{\xi\eta}(\omega)+F_{\eta\xi}(\omega), \qquad (4.15)$$

где $k^{*}(i\omega)$ — передаточная функция оптимальной линейной системы; $F_{\xi}(\omega)$, $F_{\eta}(\omega)$, $F_{\xi\eta}(\omega)$, $F_{\eta\xi}(\omega)$ — энергетические спектры и взаимные энергетические спектры процессов $\xi(t)$ и $\eta(t)$.

Итак, задача отыскания оптимальной линейной системы в рассматриваемом случае сводится к решению интегрального уравнения (4.12), которое при игнорировании физической осуществимости находится путем преобразования Фурье от обеих частей (4.12). Так как правая часть (4.12) есть свертка функций, преобразования Фурье которых равны соответственно $k^{*}(i\omega)$ и $F_x(\omega)$, то

$$F_{x\xi}(\omega)=F_{\xi}(\omega)+F_{\xi\eta}(\omega)=k^{*}(i\omega)\,F_x(\omega).$$

откуда с учетом (4.15) находим

$$k^*(i\omega)=\frac{F_{x\xi}(\omega)}{F_x(\omega)}=\frac{F_\xi(\omega)+F_{\xi\eta}(\omega)}{F_\xi(\omega)+F_\eta(\omega)+F_{\xi\eta}(\omega)+F_{\eta\xi}(\omega)}. \qquad (4.16)$$

Подставляя (4.16) в (4.14), получим величину минимальной среднеквадратической ошибки

$$\varepsilon^2_{min}=\frac{1}{2\pi}\int\limits_0^\infty \frac{F_\xi(\omega)F_x(\omega)-|F_{x\xi}(\omega)|^2}{F_x(\omega)}d\omega. \qquad (4.16')$$

Если процессы $\xi(t)$ и $\eta(t)$ некогерентны, то $F_{\xi\eta}(\omega)=F_{\eta\xi}(\omega)\equiv 0$ и из (4.16) следует

$$k^*(i\omega)=\frac{F_\xi(\omega)}{F_\xi(\omega)+F_\eta(\omega)}. \qquad (4.17)$$

Так как правая часть (4.17) действительна, то она представляет частотную характеристику оптимального линейного фильтра (фазовая характеристика в этом случае тождественно равна нулю).

Из (4.16′) при $F_{x\xi}(\omega)=F_\xi(\omega)$, $F_x(\omega)=F_\xi(\omega)+F_\eta(\omega)$ получим для некогерентных процессов

$$\varepsilon^2_{min}=\frac{1}{2\pi}\int\limits_0^\infty \frac{F_\xi(\omega)F_\eta(\omega)}{F_\xi(\omega)+F_\eta(\omega)}d\omega. \qquad (4.18)$$

Формула (4.18) показывает, что ошибка в оптимальной линейной системе может быть сделана равной нулю тогда, когда энергетические спектры процессов $\xi(t)$ и $\eta(t)$ не перекрываются, т. е. когда $F_\xi(\omega)F_\eta(\omega)\equiv 0$ при всех ω. В противном случае ошибка неизбежна. Для того чтобы не было перекрытия, необходимо, очевидно, чтобы энергетические спектры $F_\xi(\omega)$ и $F_\eta(\omega)$ на некоторых интервалах оси частот тождественно обращались в нуль *).

Таким образом, фильтрация стационарного случайного процесса из аддитивной смеси с другим стационарным случайным процессом с нулевой среднеквадратической ошибкой (сингулярность) соответствует тому случаю, когда энергетические спектры этих процессов имеют диапазоны частот, где они тождественно равны нулю.

---

*) Заметим, что один из спектров должен быть при этом финитным. Тогда их отношение равно нулю или бесконечности [см. (3.147)].

При фильтрации процесса $\xi(t)$ на фоне белого шума для которого $F_\eta(\omega) = N_0$, из (4.18) следует

$$\varepsilon^2_{min} = \frac{N_0}{2\pi} \int_0^\infty \frac{F_\xi(\omega)\, d\omega}{F_\xi(\omega) + N_0}, \tag{4.19}$$

т. е. наличие аддитивного белого шума в данном случае исключает сингулярность (см. § 3.4.6).

4.2.2. Физически реализуемая оптимальная линейная система; конечное время наблюдения. Линейная система, импульсная переходная функция которой $h^*(t, y)$ находится из уравнения (4.7), не реализуема физически, так как $h^*(t, y)$ не равна нулю при $t < y$. Полученные в предыдущем разделе соотношения можно трактовать следующим образом: операция фильтрации осуществляется после наблюдения реализации $x(t)$ суммарного процесса $\xi(t) + \eta(t)$ на всем интервале времени от $-\infty$ до ∞. Следовательно, подобная оптимальная линейная система совершает оценку значения процесса в заданный момент времени с бесконечным запаздыванием. Условие физической осуществимости означает, что для фильтрации используется только предыстория реализации $x(t)$ до момента времени, когда производится оценка, иначе говоря:

$$h(t, \tau) \equiv 0, \quad \tau > t. \tag{4.20}$$

Подставляя (4.20) в (4.1), получим оценку $\xi(t)$ при помощи физически реализуемого линейного фильтра

$$\hat{\xi}(t) = \int_{-\infty}^{t} h(t, \tau)\, x(\tau)\, d\tau. \tag{4.21}$$

В оценке (4.21) используются все значения реализации, предшествующие моменту оценки t. Если имеется реализация $x(\tau)$ конечной длительности T, т. е. если оценка в момент t производится по результатам наблюдения на замкнутом интервале $(t - T, t)$, то вместо (4.21) будем иметь

$$\hat{\xi}(t) = \int_{t-T}^{t} h(t, \tau)\, x(\tau)\, d\tau = \int_0^T h(t, t-u)\, x(t-u)\, du. \tag{4.22}$$

Подставляя (4.22) в (4.2) и выполняя те же преобразования, что и в (4.3), придем к следующему выражению среднего квадрата ошибки, учитывающего физическую реализуе-

мость фильтра и конечную длительность реализации процесса:

$$\varepsilon^2(t) = B_\xi(t, t) - 2\int_0^T h(t, t-\tau) B_{x\xi}(t-\tau, t)\, d\tau +$$
$$+ \int_0^T \int_0^T h(t, t-u)\, h(t, t-v)\, B_x(t-u, t-v)\, du\, dv. \quad (4.23)$$

Используя (4.23), нетрудно (как и в § 4.2.1) показать, что наилучшую по критерию минимума среднего квадрата ошибки линейную фильтрацию $\xi(t)$ из его аддитивной смеси с $\eta(t)$ осуществляет такая линейная система, импульсная переходная функция которой удовлетворяет интегральному уравнению

$$B_{x\xi}(t-\tau, t) = \int_0^T h^*(t, t-y)\, B_x(t-\tau, t-y)\, dy, \quad 0 \leqslant \tau \leqslant T. \quad (4.24)$$

Минимальное значение среднего квадрата ошибки, соответствующее фильтрации при помощи оптимальной линейной системы [см. (4.9′)], равно

$$\varepsilon^2_{min}(t) = B_\xi(t, t) - \int_0^T h^*(t, t-u)\, B_{x\xi}(t-u, t)\, du. \quad (4.25)$$

Для стационарных и стационарно связанных процессов $\xi(t)$ и $\eta(t)$ и для фильтра с постоянными параметрами из (4.24) и (4.25) следует:

$$B_{x\xi}(\tau) = \int_0^T h^*(y)\, B_x(\tau - y)\, dy, \quad 0 \leqslant \tau \leqslant T, \quad (4.26)$$

$$\varepsilon^2_{min} = B_\xi(0) - \int_0^T h^*(u)\, B_{x\xi}(u)\, du. \quad (4.27)$$

Если в приведенных формулах устремить T к бесконечности, то приходим к случаю бесконечного времени наблюдения, предшествующего моменту оценки.

4.2.3. Оценка линейно-преобразованного случайного процесса. Теперь распространим изложенную выше методику на оценку значения не процесса $\xi(t)$, а его линейного преобразования $\zeta(t)$. Как и раньше, оценка $\zeta(t)$ производится

линейной фильтрацией реализации $x(t)$ аддитивной смеси $\xi(t)$ и $\eta(t)$, наблюдаемой на конечном интервале времени $(t-T,\ t)$, т. е.

$$\hat{\zeta}(t)=\int_0^T h_1(t,\ t-u)\,x(t-u)\,du. \qquad (4.28)$$

Средний квадрат ошибки при использовании такой оценки равен

$$\varepsilon_\zeta^2(t)=m_1\left\{\left[\zeta(t)-\int_0^T h_1(t,\ t-u)\,x(t-u)\,du\right]^2\right\}=$$

$$=B_\zeta(t,\ t)-2\int_0^T h_1(t,\ t-\tau)\,B_{x\zeta}(t-\tau,\ t)\,d\tau+$$

$$+\int_0^T\int_0^T h_1(t,\ t-u)\,h_1(t,\ t-v)\,B_x(t-u,\ t-v)\,du\,dv, \qquad (4.29)$$

где

$$B_{x\zeta}(t_1,\ t_2)=m_1\{x(t_1)\,\zeta(t_2)\}=B_{\xi\zeta}(t_1,\ t_2)+B_{\eta\zeta}(t_1,\ t_2). \qquad (4.29')$$

Наилучшая по критерию минимума ε_ζ^2 оценка $\hat{\zeta}(t)$ осуществляется линейным фильтром, импульсная переходная функция $h_1^*(t,\ \tau)$ которого удовлетворяет интегральному уравнению [см. (4.24)]

$$B_{x\zeta}(t-\tau,\ t)=\int_0^T h_1^*(t,\ t-y)\,B_x(t-\tau,\ t-y)\,dy,\quad 0\leqslant\tau\leqslant T. \qquad (4.30)$$

При этом минимум среднего квадрата ошибки равен

$$\varepsilon_{\zeta\,min}^2=B_\zeta(t,\ t)-\int_0^T h_1^*(t,\ t-u)\,B_{x\zeta}(t-u,\ t)\,du. \qquad (4.30')$$

Для стационарных процессов формулы (4.30) и (4.30′) преобразуются к виду

$$B_{x\zeta}(\tau)=\int_0^T h_1^*(y)\,B_x(\tau-y)\,dy,\quad 0\leqslant\tau\leqslant T, \qquad (4.31)$$

$$\varepsilon_{\zeta\,min}^2=B_\zeta(0)-\int_0^T h_1^*(u)\,B_{x\zeta}(u)\,du. \qquad (4.31')$$

Различные виды линейного преобразования $\xi(t)$ отражаются лишь на выражениях корреляционной функции $B_\zeta(t_1, t_2)$ и взаимной корреляционной функции $B_{x\zeta}(t_1, t_2)$. Приведем эти выражения для трех наиболее важных видов линейных преобразований.

Пусть

$$\zeta(t) = \xi(t + t_0). \tag{4.32}$$

Тогда

$$B_\zeta(t, t) = B_\xi(t + t_0, t + t_0), \tag{4.33}$$

$$B_{x\zeta}(t_1, t_2) = B_\xi(t_1, t_2 + t_0) + B_{\eta\xi}(t_1, t_2 + t_0). \tag{4.33'}$$

Случай $t_0 > 0$ или $t_0 < -T$ соответствует *экстраполяции* случайного процесса*), а $-T < t_0 < 0$ соответствует *интерполяции*.

Если $\xi(t)$ дифференцируем в среднеквадратическом и

$$\zeta(t) = \xi'(t), \tag{4.34}$$

то (см. § 4.3.3 в первой книге)

$$B_\zeta(t) = \left.\frac{\partial^2 B_\xi(t_1, t_2)}{\partial t_1 \partial t_2}\right|_{t_1 = t_2 = t}, \tag{4.35}$$

$$B_{x\zeta}(t_1, t_2) = \frac{\partial B_\xi(t_1, t_2)}{\partial t_2} + \frac{\partial B_{\eta\xi}(t_1, t_2)}{\partial t_2}. \tag{4.35'}$$

Для процесса

$$\zeta(t) = \int_{-\infty}^{\infty} g(t, \tau)\, \xi(\tau)\, d\tau, \tag{4.36}$$

где $g(t, \tau)$ — *заданная* функция, получим (см. § 5.2.1 в первой книге)

$$B_\zeta(t, t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(t, v_1)\, g(t, v_2)\, B_\xi(v_1, v_2)\, dv_1\, dv_2, \tag{4.37}$$

$$B_{x\zeta}(t_1, t_2) = \int_{-\infty}^{\infty} g(t_2, v)\, B_\xi(t_1, v)\, dv +$$

$$+ \int_{-\infty}^{\infty} g(t_2, v)\, B_{\eta\xi}(t_1, v)\, dv. \tag{4.37'}$$

---

*) Случай $t_0 > 0$ иногда называют предсказанием.

В этом случае оценку $\hat{\zeta}(t)$ [см. (4.28)] по реализации суммы $\xi(t)+\eta(t)$, наблюдаемой на конечном интервале времени, можно трактовать как оценку, оптимальную по критерию минимума среднего квадрата ошибки, процесса на выходе линейной системы с заданной импульсной переходной функцией $g(t, \tau)$, когда на ее вход действует процесс $\xi(t)$.

Заметим, что из (4.36) — (4.37′) при $g(t, \tau) = \delta(\tau - t - t_0)$ получаются формулы (4.32) — (4.33′), а при $g(t, \tau) = \delta'(t-\tau)$ — формулы (4.34)—(4.35′).

Формулами (4.30) и (4.31) можно воспользоваться для оптимальной линейной оценки процесса $\zeta(t)$, представляющего результат преобразования $\xi(t)$ в линейной системе со случайными параметрами. Используя (5.112) из первой книги, можно функцию $B_\zeta(t, t)$ записать в виде

$$B_\zeta(t, t) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} B_c(t, t, \omega_1, \omega_2) \times$$

$$\times B_\xi(t-u, t-v)\, e^{i(\omega_1 u + \omega_2 v)}\, d\omega_1\, d\omega_2\, du\, dv, \qquad (4.38)$$

где

$$B_c(t_1, t_2, \omega_1, \omega_2) = m_1\{k(i\omega_1, t_1)\, k(i\omega_2, t_2)\} \qquad (4.38')$$

— корреляционная функция линейной системы. Предполагается, что передаточная функция линейной системы $k(i\omega, t)$ (представляющая случайный процесс) независима от $\xi(t)$.

Взаимная корреляционная функция $B_{x\zeta}(t_1, t_2)$ в этом случае записывается следующим образом:

$$B_{x\zeta}(t_1, t_2) = \frac{1}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k_c(i\omega, t_2) \times$$

$$\times [B_\xi(t_1, t_2 - v) + B_{\xi\eta}(t_1, t_2 - v)]\, e^{i\omega v}\, d\omega\, dv, \qquad (4.39)$$

где

$$k_c(i\omega, t) = m_1\{k(i\omega, t)\} \qquad (4.39')$$

— среднее значение передаточной функции линейной системы.

4.2.4. Фильтрация как задача регрессии. Рассмотренной выше задаче линейной фильтрации может быть дана не-

сколько иная трактовка. Пусть $\xi(t)$ и $\zeta(t)$ — два *зависимых* случайных процесса, и необходимо оценить значение процесса $\zeta(t)$ путем линейной фильтрации реализации $x(t)$ процесса $\xi(t)$, выбирая характеристику фильтра по критерию минимума среднего квадрата ошибки. Импульсная переходная функция такого фильтра находится из интегрального уравнения (4.30) (или (4.31) для стационарных процессов).

Таким образом сформулированная задача фильтрации представляет обобщение на случайные процессы задачи линейной среднеквадратической регрессии для случайных величин (см. приложение V). Интегральное уравнение Винера — Хопфа (4.31) аналогично системе линейных алгебраических уравнений, определяющих гиперплоскость среднеквадратической регрессии в конечномерном пространстве выборок [см. (7) в приложении V].

В общем виде функционал регрессии задает оценку процесса $\zeta(t)$ по наблюдаемой реализации $x(t)$ процесса $\xi(t)$ как *условное среднее*

$$\hat{\zeta}(t) = m_1\{\zeta(t) \mid x(t)\}.$$

Эта оценка минимизирует средний квадрат ошибки $m_1\{(\zeta - \hat{\zeta})^2\}$. Условное среднее не всегда является *линейным* функционалом реализации $x(t)$. Только в том случае, когда процессы $\xi(t)$ и $\zeta(t)$ нормальные (совместно), функционал регрессии всегда линейный (см. приложение V, а также § 4.4.4).

4.2.5. Фильтрация квазидетерминированного сигнала. В качестве первого примера, иллюстрирующего теорию оптимальной фильтрации, рассмотрим задачу о фильтрации квазидетерминированного сигнала $\xi(t) = as(t)$, вид которого определяется известной функцией $s(t)$, а амплитуда a случайна. Сигнал наблюдается в аддитивной смеси с независящим от a случайным процессом $\eta(t)$, среднее значение которого равно нулю, а корреляционная функция равна $B(t_1, t_2)$. Необходимо найти характеристику линейного фильтра, который формирует наилучшим образом (по критерию минимума среднего квадрата ошибки оценку $\hat{\xi}(t) = \hat{a}\, s(t)$, по реализации аддитивной смеси $\xi(t) + \eta(t)$, наблюдаемой на интервале $(t - T, t)$. Для этого необходимо решить интегральное уравнение (4.24), подставив в него вместо $B_{x\xi}$ и B_x следующие выражения

[см. (4.4) и (4.5)]:

$$B_{x\xi}(t-\tau,\ t)=m_1\{a^2\}\,s(t-\tau)\,s(t), \qquad (4.40)$$

$$B_x(t-\tau,\ t-y)=m_1\{a^2\}\,s(t-\tau)\,s(t-y)+B(t-\tau,\ t-y). \qquad (4.40')$$

Вводя обозначение для второго момента амплитуды

$$m^2=m_1\{a^2\},$$

получим после указанной подстановки

$$m^2 s(t-\tau)\,s(t)=\int_0^T h^*(t,\ t-y)\,[m^2 s(t-\tau)\,s(t-y)+B(t-\tau,\ t-y)]\,dy, \quad 0\leqslant\tau\leqslant T. \qquad (4.41)$$

Решение интегрального уравнения (4.41) ищем в виде

$$h^*(t,\ \tau)=\lambda s(t)\,v(\tau), \qquad (4.42)$$

где $v(\tau)$ — решение другого интегрального уравнения *)

$$\int_{t-T}^{t} v(u)\,B(y,\ u)\,du=s(y), \quad t-T\leqslant y\leqslant t. \qquad (4.43)$$

Подставляя (4.42) в (4.41), с учетом (4.43) получаем

$$m^2=\lambda\left[m^2\int_0^T v(t-y)\,s(t-y)\,dy+1\right],$$

откуда находим величину λ:

$$\lambda=m^2\left(1+m^2\int_{t-T}^{t} v(z)\,s(z)\,dz\right)^{-1}. \qquad (4.44)$$

Итак, оптимальная фильтрация квазидетерминированного сигнала $as(t)$ на фоне случайного процесса $\eta(t)$ осуществляется линейной системой с импульсной переходной функцией

$$h^*(t,\ \tau)=\frac{m^2 s(t)\,v(\tau)}{1+m^2\int_{t-T}^{t} v(z)\,s(z)\,dz}, \qquad (4.45)$$

где $v(\tau)$ — решение интегрального уравнения (4.43). Оценка сигнала в момент t по реализации $x(\tau)$ его аддитивной

*) Уравнение типа (4.43) уже неоднократно появлялось в предыдущей главе.

смеси с $\eta(t)$, наблюдаемой на интервале $(t-T,\ t)$, имеет вид

$$\hat{a}s(t)=\int_0^T h^*(t,\ t-u)\,x(t-u)\,du=$$

$$=\frac{m^2 s(t)\int\limits_{t-T}^{t} v(z)\,x(z)\,dz}{1+m^2\int\limits_{t-T}^{t} v(z)\,s(z)\,dz}, \tag{4.46}$$

а оценка случайной амплитуды сигнала *)

$$\hat{a}=\frac{m^2\int\limits_{t-T}^{t} v(z)\,x(z)\,dz}{1+m^2\int\limits_{t-T}^{t} v(z)\,s(z)\,dz}. \tag{4.46'}$$

Оценка $\hat{a}$ *смещенная*, так как из (4.46′) следует:

$$m_1\{\hat{a}\}=m_1\{a\}\frac{m^2\int\limits_{t-T}^{t} v(z)\,s(z)\,dz}{1+m^2\int\limits_{t-T}^{t} v(z)\,s(z)\,dz}, \tag{4.47}$$

но, как видно из (4.47), эта оценка асимптотически при $m^2\to\infty$ несмещенная.

Найдем средний квадрат ошибки при использовании оценки (4.46). Согласно (4.30′) и (4.40′) имеем

$$\varepsilon^2_{min}(t)=m^2s^2(t)-\lambda m^2\int_0^T s^2(t)\,s(t-\tau)\,v(t-\tau)\,d\tau=$$

$$=m^2s^2(t)\left[1-\lambda\int\limits_{t-T}^{t} v(z)\,s(z)\,dz\right]$$

---

*) Подчеркнем, что оценка (4.46′) представляет оптимальную по критерию среднего квадрата ошибки оценку в классе линейных функционалов от наблюдаемой реализации $x(t)$. При этом не делается *никаких* предположений о функции распределения аддитивного шума. При $m\to\infty$ эта оценка совпадает с оценкой максимального правдоподобия амплитуды детерминированного сигнала на фоне аддитивного *нормального* шума [см. (3.166)]. Заметим, однако, что (4.46′) совпадает с байесовской оценкой (3.190) при квадратичной функции потерь и нормальном распределении шума, если в (3.190) положить $\sigma_0^2=m^2$ и $a_0=0$.

или с учетом (4.44)

$$\varepsilon^2_{min}(t) = m_1\{(\hat{a} - a)^2 s^2(t)\} = \\ = \frac{m^2 s^2(t)}{1 + m^2 \int\limits_{t-T}^{t} v(z)\, s(z)\, dz}. \tag{4.48}$$

Из (4.48) следует также, что

$$m_1\{(\hat{a} - a)^2\} = \frac{m^2}{1 + m^2 \int\limits_{t-T}^{t} v(z)\, s(z)\, dz}, \tag{4.49}$$

причем при $m^2 \to \infty$ средний квадрат отклонения оценки $\hat{a}$ от оцениваемой амплитуды a сигнала стремится к своему минимальному значению, равному $\left(\int\limits_{t-T}^{t} v(z)\, s(z)\, dz\right)^{-1}$.

Отношение

$$\frac{m_1\{(\hat{a} - a)^2\}}{m_1\{(\hat{a} - a)^2\}_{m^2=\infty}} = \mathrm{e} = \frac{m^2 \int\limits_{t-T}^{t} v(z)\, s(z)\, dz}{1 + m^2 \int\limits_{t-T}^{t} v(z)\, s(z)\, dz}, \tag{4.50}$$

которое не превосходит единицы, указывает, насколько средний квадрат ошибки при конечном втором моменте амплитуды меньше этой величины при бесконечном втором моменте. Сравнивая (4.50) с (4.47), замечаем, что

$$m_1\{\hat{a}\} = \mathrm{e}\, m_1\{a\}. \tag{4.50'}$$

В частном случае, когда фильтрация квазидетерминированного сигнала происходит на фоне белого шума со спектральной интенсивностью N_0,

$$B(t, u) = N_0 \delta(t - u), \quad v(\tau) = \frac{1}{N_0} s(\tau),$$

и из (4.45), (4.48) получаем

$$h^*(t, \tau) = \frac{\frac{m^2}{N_0} s(t)\, s(\tau)}{1 + \frac{m^2}{N_0} \int\limits_{t-T}^{t} s^2(z)\, dz}, \tag{4.51}$$

$$\varepsilon^2_{min}(t) = \frac{m^2 s^2(t)}{1 + \frac{m^2}{N_0} \int\limits_{t-T}^{t} s^2(z)\, dz}. \tag{4.52}$$

Приведенные выше соотношения остаются, конечно, справедливыми и для случая фильтрации детерминированного сигнала, если под величиной m понимать амплитуду детерминированного сигнала. Все результаты, относящиеся к оценке неизвестной (не случайной) амплитуды детерминированного сигнала, указанные в § 3.6.1, получают из соответствующих формул этого раздела при $m^2 \to \infty$ (см. подстрочное примечание на стр. 293).

Заметим, что в соответствии с (4.46) *дисперсия* оценки квазидетерминированного сигнала по реализации аддитивной смеси сигнала и шума равна дисперсии шумового слагаемого на выходе фильтра, т. е.

$$M_2\{\hat{as}(t)\} = \int_0^T \int_0^T h^*(t,\ t-u)\, h^*(t,\ t-v) \times$$

$$\times B(t-u,\ t-v)\, du\, dv = \lambda^2 s^2(t) \int_{t-T}^{t} v(z)\, s(z)\, dz =$$

$$= m^2 s^2(t) \frac{m^2 \int\limits_{t-T}^{t} v(z)\, s(z)\, dz}{\left(1 + m^2 \int\limits_{t-T}^{t} v(z)\, s(z)\, dz\right)^2}. \qquad (4.53)$$

Сравнивая (4.53) с (4.48), находим

$$\frac{M_2\{\hat{as}(t)\}}{\varepsilon^2_{min}(t)} = \frac{m^2 \int\limits_{t-T}^{t} v(z)\, s(z)\, dz}{1 + m^2 \int\limits_{t-T}^{t} v(z)\, s(z)\, dz} < 1, \qquad (4.54)$$

причем это отношение стремится к единице, когда $m^2 \to \infty$.

4.2.6. Чистая экстраполяция. В качестве второго иллюстративного примера рассмотрим задачу об экстраполяции значений случайного процесса $\xi(t)$ по наблюдаемой на конечном отрезке времени реализации *этого же процесса*. Решение этой частной задачи получается из общего решения, приведенного в § 4.2.3, если положить $\zeta(t) = \xi(t + t_0)$ и $\eta(t) \equiv 0$. Экстраполированное значение $\hat{\xi}(t + t_0)$ получается линейной фильтрацией наблюдаемой реализа-

ции $x(t)$ процесса $\xi(t)$

$$\hat{\xi}(t+t_0)=\int_0^T h(t,\, t-u)\, x(t-u)\, du, \qquad (4.55)$$

а импульсная переходная характеристика наилучшего фильтра по критерию минимума среднего квадрата ошибки находится из интегрального уравнения [см. (4.30)]

$$B_\xi(t-\tau,\, t+t_0)=\int_0^T h^*(t,\, t-y)\times$$
$$\times B_\xi(t-\tau,\, t-y)\, dy, \quad 0\leqslant\tau\leqslant T. \qquad (4.56)$$

Соответствующее (минимальное) значение среднего квадрата ошибки экстраполяции равно [см. (4.30′)]

$$\varepsilon^2(t,\, t_0)=B_\xi(t+t_0,\, t+t_0)-$$
$$-\int_0^T h^*(t,\, t-u)\, B_\xi(t-u,\, t+t_0)\, du. \qquad (4.57)$$

Импульсная переходная характеристика $h^*(\tau)$ оптимальной линейной системы с постоянными параметрами для получения экстраполированного значения стационарного (в широком смысле) процесса

$$\hat{\xi}(t+t_0)=\int_0^T h^*(u)\, x(t-u)\, du \qquad (4.58)$$

является решением интегрального уравнения

$$B_\xi(\tau+t_0)=\int_0^T h^*(y)\, B_\xi(\tau-y)\, dy, \quad 0\leqslant\tau\leqslant T, \qquad (4.59)$$

которое является частным видом уравнения (4.56). Средний квадрат ошибки при этом равен

$$\varepsilon^2_{min}(t_0)=B_\xi(0)-\int_0^T h^*(u)\, B_\xi(u+t_0)\, du. \qquad (4.60)^*)$$

---

*) Заметим, что для чистой экстраполяции, без учета физической реализуемости фильтра, интегрирование в (4.58) — (4.60) производится по всем действительным значениям переменной. Интегральное уравнение при этом имеет тривиальное решение $h^*(y)=\delta(y+t_0)$; оценка совпадает со значением реализации $\hat{\xi}(t+t_0)=x(t+t_0)$, а средний квадрат ошибки равен нулю.

4.3. СОГЛАСОВАННЫЕ ФИЛЬТРЫ

4.3.1. Линейная фильтрация по критерию максимума отношения сигнал/шум. В предыдущем разделе качество линейной фильтрации определялось величиной среднего квадрата отклонения оценки от значения оцениваемого случайного процесса. Такой подход оправдан в ситуациях, когда при выделении сигнала имеет значение вид его функциональной зависимости от времени. В других радиотехнических задачах, например в радиолокации, важно установить наличие или отсутствие сигнала, скрытого в шумах, т. е. важно при фильтрации обеспечить максимальное отношение сигнал/шум даже ценой существенного искажения формы сигнала.

Пусть $x(t)$ — реализация *аддитивной* смеси детерминированного сигнала $s(t)$ и стационарного (в широком смысле) случайного шума $\eta(t)$ с нулевым средним и известной корреляционной функцией $B_\eta(\tau)$. Реализация наблюдается в течение конечного интервала времени T. Оценку сигнала дает профильтрованное значение реализации

$$\hat{s}(t) = \int_0^T h(\tau)\, x(t-\tau)\, d\tau, \tag{4.61}$$

где $h(\tau)$ — подлежащая определению импульсная переходная функция линейной системы (с постоянными во времени параметрами). Критерием качества оценки принимаем максимум отношения сигнал/шум на выходе фильтра, которое нетрудно определить в силу линейности фильтра (см. § 6.3.1 в первой книге). Действительно, (4.61) можно переписать в виде

$$\hat{s}(t) = s_1(t) + \nu(t), \tag{4.62}$$

где

$$s_1(t) = \int_0^T h(\tau)\, s(t-\tau)\, d\tau, \tag{4.62'}$$

$$\nu(t) = \int_0^T h(\tau)\, \eta(t-\tau)\, d\tau. \tag{4.62''}$$

Отношение сигнал/шум определяется как отношение квадрата сигнала на выходе $s_1^2(t_0)$ в некоторый момент време-

ни t_0 к дисперсии шума на выходе фильтра $\sigma_\nu^2 = M_2\{\nu(t)\}$, т. е.

$$\frac{\text{С}}{\text{Ш}} = \frac{s_1^2(t_0)}{\sigma_\nu^2}. \tag{4.63}$$

Оптимальный по критерию максимума отношения сигнал/шум линейный фильтр называют *согласованным*.

4.3.2. Импульсная переходная и передаточная функции согласованного фильтра. Определим импульсную переходную функцию линейной системы, для которой отношение $\frac{\text{С}}{\text{Ш}}$ имеет максимальную величину. Обозначим эту величину через μ_{max}. Тогда для любой линейной системы с импульсной переходной функцией $h(\tau)$ имеет место неравенство [см. (4.62), (4.63)]

$$\mu_{max}\sigma_\nu^2 - s_1^2(t_0) = \mu_{max}\int_0^T\int_0^T h(u)\,h(v)\,B_\eta(v-u)\,du\,dv - \left(\int_0^T h(\tau)\,s(t_0-\tau)\,d\tau\right)^2 \geqslant 0, \tag{4.64}$$

причем знак равенства соответствует согласованному фильтру. Нетрудно проверить, что левая часть (4.64) обращается в нуль для функции $h(\tau) = h^*(\tau;\ t_0)$, удовлетворяющей интегральному уравнению

$$\int_0^T h^*(\tau;\ t_0)\,B_\eta(t-\tau)\,d\tau = \mathrm{k}(t_0)\,s(t_0-t), \qquad 0 \leqslant t \leqslant T, \tag{4.65}$$

где константа $\mathrm{k}(t_0)$ равна

$$\mathrm{k}(t_0) = \frac{\int_0^T h^*(\tau;\ t_0)\,s(t_0-\tau)\,d\tau}{\mu_{max}} = \frac{s_1^*(t_0)}{\mu_{max}} = \frac{\sigma_\nu^{*2}}{s_1^*(t_0)} \tag{4.66}$$

(звездочка указывает, что значения сигнала s_1 и дисперсии шума относятся к процессу на выходе согласованного фильтра).

Заметим, что если $h^*(\tau;\ t_0)$ представляет решение уравнения (4.65), то этому уравнению удовлетворяет также функция $ch^*(\tau;\ t_0)$, где c — произвольная константа, а величина μ_{max} остается неизменной. Таким образом,

импульсную переходную функцию согласованного фильтра достаточно определить с точностью до постоянного (масштабного) множителя.

При фильтрации детерминированного сигнала из аддитивной смеси с белым шумом импульсная переходная функция согласованного фильтра может быть выражена

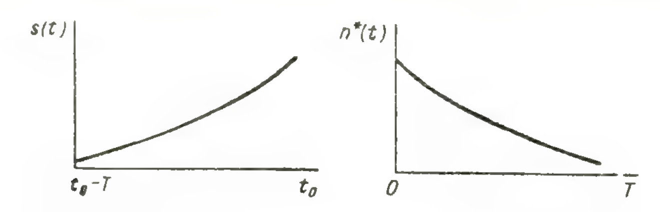


Рис. 22. Сигнал и импульсная переходная функция согласованного фильтра.

в явном виде, если в (4.65) подставить $B_\eta(t-\tau) = N_0\delta(t-\tau)$. Используя фильтрующее свойство дельта-функции, получаем

$$h^*(t;\, t_0) = \frac{k(t_0)}{N_0}\, s(t_0 - t), \quad 0 < t \leqslant T. \tag{4.67}$$

Следовательно, в рассматриваемом случае импульсная переходная функция согласованного фильтра пропорциональна зеркальному отображению относительно вертикальной оси, проходящей через точку $t = t_0$, отрезка сигнала $s(t)$ на интервале $(t_0 - T,\ t_0)$ с последующим переносом начала координат в точку $t = t_0$ (рис. 22). В том случае, когда $t_0 = T$, импульсная переходная функция получается одним зеркальным отображением сигнала относительно вертикали, делящей пополам интервал $(0,\ T)$.

Максимальное значение отношения сигнал/шум равно [см. (4.64)]

$$\mu_{max} = \frac{\left[\int\limits_0^T s^2(t_0-\tau)\,d\tau\right]^2}{N_0 \int\limits_0^T s^2(t_0-\tau)\,d\tau} = \frac{1}{N_0}\int\limits_0^T s^2(t_0-\tau)\,d\tau, \tag{4.68}$$

т. е. отношению энергии сигнала на интервале $(t_0 - T,\ t_0)$ к спектральной плотности N_0 шума.

Передаточная функция $k^*(i\omega;\ t_0)$ согласованного фильтра с импульсной переходной функцией (4.67) равна

$$k^*(i\omega;\ t_0) = \int\limits_{-\infty}^{\infty} h^*(t;\ t_0)\, e^{-i\omega t}\, dt =$$

$$= \frac{k(t_0)}{N_0} \int\limits_{0}^{T} s(t_0 - t)\, e^{-i\omega t}\, dt =$$

$$= \frac{k(t_0)}{N_0}\, e^{-i\omega t_0} \int\limits_{t_0 - T}^{t_0} s(t)\, e^{i\omega t}\, dt$$

или

$$k^*(i\omega;\ t_0) = \frac{k(t_0)}{N_0}\, \overline{A_s(\omega;\ t_0,\ T)}\, e^{-i\omega t_0}, \qquad (4.69)$$

где $A_s(\omega;\ t_0,\ T)$ — спектр отрезка сигнала $s(t)$ на интервале $(t_0 - T,\ t_0)$ и черта сверху указывает на комплексно-сопряженную величину.

Итак, при выделении сигнала из аддитивной смеси с белым шумом передаточная функция согласованного фильтра пропорциональна комплексно-сопряженному спектру усеченного сигнала.

Можно получить общее выражение передаточной функции согласованного фильтра при произвольном энергетическом спектре $F_\eta(\omega)$ шума, если, отказавшись от условия физической осуществимости, предположить, что имеется реализация $x(t)$ для всех действительных t, и представить оценку сигнала в виде [ср. с (4.61)]

$$\hat{s}(t) = \int\limits_{-\infty}^{\infty} h^*(\tau;\ t_0)\, x(t - \tau)\, d\tau.$$

В этом случае уравнение (4.65) преобразуется к виду (для всех действительных значений t)

$$\int\limits_{-\infty}^{\infty} h^*(\tau;\ t_0)\, B_\eta(t - \tau)\, d\tau = k(t_0)\, s(t_0 - t). \qquad (4.70)$$

Совершая преобразование Фурье над обеими частями (4.70), приходим к выражению

$$k^*(i\omega;\ t_0)\, F_\eta(\omega) = k(t_0)\, \overline{A_s(\omega)}\, e^{-i\omega t_0}$$

(здесь $A_s(\omega)$ – спектр *неусеченного* сигнала), из которого находим передаточную функцию согласованного фильтра:

$$k^*(i\omega;\ t_0) = \mathrm{k}(t_0)\frac{\overline{A_s(\omega)}}{F_\eta(\omega)}\,\mathrm{e}^{-i\omega t_0}. \qquad (4.71)$$

Максимальное отношение сигнал/шум в соответствии с (4.64) равно (см. также § 5.2.1 в первой книге)

$$\mu_{max} = \frac{\left[\int\limits_{-\infty}^{\infty} h^*(\tau,\ t_0)\, s(t_0-\tau)\, d\tau\right]^2}{\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} h^*(u;\ t_0)\, h^*(v;\ t_0)\, B_\eta(v-u)\, du\, dv} =$$

$$= \frac{\left[\frac{1}{2\pi}\int\limits_{-\infty}^{\infty} k^*(i\omega;\ t_0)\, A_s(\omega)\,\mathrm{e}^{i\omega t}\, d\omega\right]^2}{\frac{1}{2\pi}\int\limits_{-\infty}^{\infty} F_\eta(\omega)\,|\,k^*(i\omega;\ t_0)\,|^2\, d\omega},$$

и, подставляя вместо $k^*(i\omega;\ t_0)$ его выражение из (4.71), получаем окончательно

$$\mu_{max} = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} \frac{|\,A_s(\omega)\,|^2}{F_\eta(\omega)}\, d\omega. \qquad (4.72)$$

В случае белого шума $F_\eta(\omega) = N_0$ и из (4.72) следует [ср. с (4.68)]:

$$\mu_{max} = \frac{1}{2\pi N_0}\int\limits_{-\infty}^{\infty} |\,A_s(\omega)\,|^2\, d\omega. \qquad (4.72')$$

Отметим, что оценка сигнала на выходе согласованного фильтра, вообще говоря, *смещенная*, так как

$$m_1\{\hat{s}(t)\} = m_1\left\{\int\limits_0^T h^*(\tau;\ t_0)\, x(t-\tau)\, d\tau\right\} =$$

$$= \int\limits_0^T h^*(\tau;\ t_0)\, s(t-\tau)\, d\tau = s_1^*(t) \neq s(t). \qquad (4.73)$$

Дисперсия этой оценки равна дисперии шумового слагаемого $\nu(t)$ на выходе согласованного фильтра [см. (4.53)]

$$M_2\{\hat{s}(t)\} = \sigma_\nu^{*2} \int_0^T \int_0^T h^*(u;\, t_0)\, h^*(v;\, t_0)\, B_\eta(v-u)\, du\, dv =$$

$$= \mathrm{k}(t_0) \int_0^T h^*(u;\, t_0)\, s(t_0-u)\, du =$$

$$= \mathrm{k}(t_0)\, s_1^*(t_0) = \frac{[s_1^*(t_0)]^2}{\mu_{max}}. \qquad (4.73')$$

В случае белого шума из (4.73), (4.73′) и (4.67) следует:

$$m_1\{\hat{s}(t)\} = \frac{\mathrm{k}(t_0)}{N_0} \int_0^T s(t_0-\tau)\, s(t-\tau)\, d\tau,$$

$$M_2\{\hat{s}(t)\} = \frac{\mathrm{k}^2(t_0)}{N_0} \int_0^T s^2(t_0-\tau)\, d\tau.$$

4.3.3. Оптимальная фильтрация периодической последовательности импульсов из аддитивной смеси с белым шумом. В качестве простейшего примера, иллюстрирующего результаты предыдущего раздела, найдем частотную характеристику согласованного фильтра, предназначенного для выделения периодической последовательности импульсов из аддитивного белого шума. Пусть τ_0 — длительность одиночного импульса и пусть $u(t)$ — функция, отличная от нуля на интервале $0 \leqslant t \leqslant \tau_0$, описывает форму этого импульса. Спектр импульса имеет вид

$$A_u(\omega) = \int_0^{\tau_0} u(t)\, \mathrm{e}^{-i\omega t}\, dt.$$

Для периодической последовательности импульсов с периодом T

$$s(t) = \sum_{n=-\infty}^{\infty} u(t-nT)$$

спектр равен

$$A_s(\omega) = \sum_{n=-\infty}^{\infty} A_u(\omega)\, \mathrm{e}^{-in\omega T} = \frac{A_u(\omega)}{1-\mathrm{e}^{-i\omega T}}.$$

Нормированная частотная характеристика согласованного фильтра (т. е. отношение модуля коэффициента пере-

дачи к $\frac{k(t_0)}{N_0}$) будет согласно (4.71) иметь следующий вид:

$$C(\omega) = |A_s(\omega)| = \frac{|A_u(\omega)|}{2\left|\sin\frac{\omega T}{2}\right|}. \tag{4.74}$$

Если длительность одиночного импульса стремится к нулю, то $|A_u(\omega)|$ остается постоянным в увеличивающемся диа-

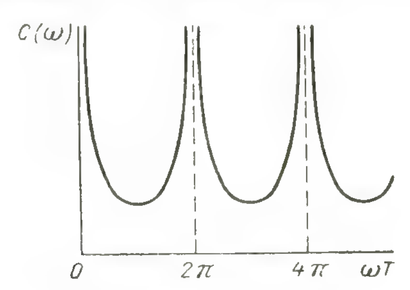


Рис. 23. Частотная характеристика гребенчатого фильтра.

пазоне частот и частотная характеристика согласованного фильтра приближается к частотной характеристике идеального гребенчатого фильтра (рис. 23):

$$C(\omega) \sim \frac{1}{2\left|\sin\frac{\omega T}{2}\right|}. \tag{4.74'}$$

Если сигнал представляет пачку импульсов конечной длительности (с числом импульсов N), то

$$A_s(\omega; N) = \sum_{n=0}^{N-1} A_u(\omega)\, e^{-in\omega T} = A_u(\omega)\frac{1-e^{-iN\omega T}}{1-e^{-i\omega T}}$$

и в соответствии с (4.69) нормированная частотная характеристика согласованного фильтра определяется по формуле

$$C(\omega) = |A_u(\omega)|\left|\frac{\sin\frac{N\omega T}{2}}{\sin\frac{\omega T}{2}}\right|. \tag{4.75}$$

Заметим, что в отличие от частотной характеристики (4.74), которая не ограничена на частотах $\frac{2\pi k}{T}$, где k –

любое целое и $\left| A_u \left(\frac{2\pi k}{T} \right) \right| > 0$, частотная характеристика согласованного фильтра для импульсной последовательности конечной длительности ограничена и на частотах, кратных тактовой частоте, так как из (4.75) следует, что

$$C\left(\frac{2\pi k}{T}\right) = N \left| A_u \left(\frac{2\pi k}{T}\right) \right|.$$

4.3.4. Активный и пассивный фильтры. Устройство, вырабатывающее оптимальную по критерию максимума отношения сигнал/шум оценку сигнала

$$\hat{s}(t) = \int_0^T h^*(\tau)\, x(t-\tau)\, d\tau \tag{4.76}$$

(где $h^*(\tau)$ – решение неоднородного интегрального уравнения (4.76) и $x(t)$ — наблюдаемая на интервале $(t - T, t)$ аддитивная смесь сигнала и шума), может быть интерпретировано двумя способами. Согласно одному из этих способов, указанному выше, согласованный фильтр, выделяющий сигнал из аддитивной смеси со стационарным случайным шумом, интерпретируется как *линейная система* с постоянными параметрами, импульсная переходная функция $h^*(\tau)$ которой определяется видом сигнала и корреляционной функцией шума [см. (4.65)]. В частном случае белого шума эта импульсная переходная функция получается зеркальным отображением сигнала.

Однако устройство, на выходе которого получается оценка (4.76), можно интерпретировать иначе. Пусть имеется генератор, вырабатывающий на основании решения интегрального уравнения (4.65) функцию $h^*(\tau)$. Тогда оценка $\hat{s}(t)$ получается при помощи устройства типа *коррелометра*, в котором функция $h^*(\tau)$ от местного генератора перемножается с принимаемой реализацией, прошедшей линию задержки, а полученное произведение интегрируется по всему интервалу наблюдения. Подобное корреляционное устройство для выделения сигнала иногда называют *активным фильтром* в отличие от линейной системы с импульсной переходной функцией $h^*(\tau)$, которую называют *пассивным фильтром* *).

*) По терминологии Д. Миддлтона, активный и пассивный фильтры представляют согласованные фильтры первого рода первого и второго типа соответственно [10]

4.4. НЕЛИНЕЙНАЯ ФИЛЬТРАЦИЯ ПО КРИТЕРИЮ МИНИМУМА СРЕДНЕГО КВАДРАТА ОШИБКИ

4.4.1. Общий метод характеризации нелинейных систем. В предыдущих разделах рассматривались оценки случайного процесса, получаемые линейной фильтрацией наблюдаемой реализации, и определялись характеристики линейных фильтров, оптимальных в смысле принятого критерия качества оценки. Если отказаться от условия линейности алгоритма обработки наблюдаемой реализации, то в более широком классе допускаемых оценок могут быть, вообще говоря, получены оценки, которые по заданному критерию будут лучше линейных оценок.

При использовании линейной системы как устройства получения оценки связь между оценкой и наблюдаемой реализацией дается достаточно простым интегральным соотношением (4.61). Если для оценки используется нелинейная инерционная система (с возможными обратными связями), то такая простая связь оценки с наблюдаемой реализацией уже не имеет места. Если определить нелинейную систему некоторым нелинейным интегродифференциальным уравнением, то при этом часто возникают значительные математические трудности. Можно, однако, попытаться и для нелинейной системы установить явную зависимость процесса на ее выходе от процесса на входе.

Пример такой зависимости дает формула (9.5) в первой книге, которая выражает процесс $\zeta(t)$ на выходе типового звена (усилитель — квадратичный детектор — фильтр) через процесс $x(t)=s(t)+\xi(t)$ на его входе в виде двойного интеграла

$$\zeta(t)=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}K(u, v)\,x(t-u)\,x(t-v)\,du\,dv, \qquad (4.77)$$

где

$$K(u, v)=\int\limits_{-\infty}^{\infty}h_1(u-\tau)\,h_2(\tau)\,h_1(v-\tau)\,d\tau; \qquad (4.78)$$

$h_1(\tau)$, $h_2(\tau)$ — импульсные переходные функции усилителя и фильтра соответственно.

Заметим, что для нелинейной системы усилитель — квадратичный детектор, т. е. при $h_2(\tau)=\delta(\tau)$, ядро

$K(u, v)$ равно произведению

$$K(u, v) = h_1(u)\, h_1(v). \tag{4.78'}$$

Представление вида (4.77) можно обобщить на тот случай, когда характеристика нелинейного неинерционного элемента типового звена необязательно параболическая, а аппроксимируется конечным числом членов степенного ряда

$$z = \sum_{k=1}^{n} a_k y^k.$$

Тогда вместо (4.77) получим

$$\zeta(t) = \sum_{k=1}^{n} a_k \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} K(u_1, \ldots, u_k)\, x(t-u_1) \ldots \\ \ldots x(t-u_k)\, du_1 \ldots du_k, \tag{4.79}$$

где

$$K(u_1, \ldots, u_k) = \int_{-\infty}^{\infty} h_1(u_1-\tau) \ldots h_1(u_k-\tau)\, h_2(\tau)\, d\tau. \tag{4.80}$$

При $h_2(\tau) = \delta(\tau)$ многомерное ядро сводится к произведению функций одного переменного

$$K(u_1, \ldots, u_k) = \prod_{i=1}^{k} h_1(u_i). \tag{4.80'}$$

Хотя во многих случаях сложные схемы нелинейных систем, содержащие обратные связи, могут быть заменены эквивалентными многокаскадными типовыми звеньями, укажем на общий подход к вопросу о том, как охарактеризовать нелинейную систему произвольного вида. Отправным пунктом этого подхода является вполне естественное утверждение о том, что процесс на выходе нелинейной системы представляет функционал, заданный на множестве возможных процессов, действующих на ее вход. Далее используется одна замечательная теорема, доказанная еще в 1910 г. французским математиком Фреше *). Он показал, что для любого непрерывного функционала **) $y[x(t)]$ существует последовательность функционалов $y_n[x(t)]$, которая при $n \to \infty$ сколь угодно точно аппроксимирует $y[x(t)]$.

---

*) Frechet M. Sur le functionelles continues. Ann. de l'Ecole Normal Sup., 3-rd Ser., 1910, p. 27.

**) О понятии непрерывности в функциональном пространстве см., например, учебник: Шилов Г. Е. Математический анализ. Специальный курс. Физматгиз, 1961, гл. 2, § 7.

Аппроксимирующая последовательность представляется суммой интегралов (так называемых интегралов Вольтерра)

$$y_n[x(t)] = K_0 + \int_{-\infty}^{\infty} K_1(u_1)\, x(t-u_1)\, du_1 + $$
$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2(u_1, u_2)\, x(t-u_1)\, x(t-u_2)\, du_1\, du_2 + \ldots +$$
$$+ \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} K_n(u_1, \ldots, u_n)\, x(t-u_1) \ldots x(t-u_n)\, du_1 \ldots du_n. \quad (4.81)$$

Утверждение Фреше представляет собой теорему существования, так как в формуле (4.81) ядра $K_m(u_1, \ldots, u_m)$, $m = 1, 2, \ldots$ остаются неопределенными, а гарантируется лишь возможность в каждом конкретном случае найти последовательность аппроксимирующих функционалов.

Теорема Фреше аналогична известной теореме Вейерштрасса об аппроксимации непрерывной функции последовательностью полиномов, которая также относится к теоремам существования. Практически для аппроксимации непрерывных функций используется та или иная совокупность ортогональных полиномов. Подобными методами могут быть определены соответствующие последовательности ядер $K_m(u_1, \ldots, u_m)$ для аппроксимации функционалов по формуле (4.81) (см. § 4.4.5).

Заметим, что выражение (4.79) процесса на выходе типового звена является частным случаем (4.81) при

$$K_m(u_1, \ldots, u_m) = K(u_1, \ldots, u_m),$$

где функции $K(u_1, \ldots, u_m)$ определены согласно (4.80).

Если $K_m \equiv 0$ для всех $m > 1$, то получаем чисто линейное преобразование $x(t)$, причем $K_1(u_1)$ — импульсная переходная функция линейного фильтра. Добавление членов ряда (4.81) при $m > 1$ означает введение нелинейности. Совокупность ядер $K_0, K_1, \ldots, K_n$ характеризует *нелинейный фильтр* n-го порядка. Линейные фильтры по этой терминологии являются фильтрами первого порядка.

Наконец, следует оговорить, что теорема Фреше относится, строго говоря, к функционалам, определенным на множестве детерминированных непрерывных функций. Будем предполагать, не вдаваясь в математические детали, что выполнены необходимые ограничения, при которых

эта теорема справедлива и при условии, что процесс на входе нелинейной системы случайный.

4.4.2. Фильтры второго порядка. Пусть $x(t)$ — реализация суммы двух стационарных и стационарно связанных случайных процессов $\xi(t)+\eta(t)$, определенная для всех действительных значений t. Предполагая, что средние значения указанных процессов равны нулю, примем в качестве оценки

$$\hat{\xi}(t)=\int_{-\infty}^{\infty} K_1(u_1)\,x(t-u_1)\,du_1+$$

$$+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2(u_1,u_2)\,x(t-u_1)\,x(t-u_2)\,du_1\,du_2, \qquad (4.82)$$

где $K_1(u_1)=h^*(u_1)$ — импульсная переходная функция оптимальной линейной системы, определенная из уравнения (4.12).

Оценка (4.82) отличается от рассмотренных выше в разделе 4.2 наличием *нелинейного слагаемого*. К оценке, оптимальной в классе линейных систем, добавляется корректирующее слагаемое за счет использования нелинейности. Для формирования оценки (4.82) использована простейшая нелинейная система — фильтр второго порядка. Задача состоит в том, чтобы определить характеристику $K_2(u_1, u_2)$ нелинейности таким образом, чтобы средний квадрат ошибки

$$\varepsilon_2^2=m_1\{[\xi(t)-\hat{\xi}(t)]^2\} \qquad (4.83)$$

был минимальным.

Обозначим через $\xi_1(t)$ ошибку, которая получается, если для оценки $\xi(t)$ используется только оптимальная линейная система, т. е.

$$\xi_1(t)=\xi(t)-\int_{-\infty}^{\infty} h^*(u_1)\,x(t-u_1)\,du_1. \qquad (4.84)$$

Заметим, что ошибка $\xi_1(t)$ *не коррелирована* с $x(t)$, так как с учетом (4.12)

$$m_1\{x(t-v)\,\xi_1(t)\}=m_1\{x(t-v)\,\xi(t)\}-$$

$$-\int_{-\infty}^{\infty} h^*(u_1)\,m_1\{x(t-v)\,x(t-u_1)\}\,du_1=$$

$$-B_{x_{\xi}}(v) - \int_{-\infty}^{\infty} h^{*}(u_1)\, B_x(u_1 - v)\, du_1 = 0,$$

т. е.

$$m_1 \{x(t - v)\, \xi_1(t)\} = 0. \tag{4.85}$$

Из (4.82) — (4.84) с учетом стационарности и стационарной связанности процессов находим

$$\varepsilon_2^2 = B_{\xi_1}(0) - 2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K_2(u_1, u_2)\, m_{\xi_1 x}(u_1, u_2)\, du_1\, du_2 +$$

$$+ \int_{\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{\infty}^{\infty} K_2(u_1, u_2) \times$$

$$\times K_2(u_3, u_4)\, m_x(u_1 - u_2,\ u_1 - u_3,\ u_2 - u_4)\, du_1\, du_2\, du_3\, du_4, \tag{4.86}$$

где

$$m_{\xi_1 x}(u_1, u_2) = m_1 \{\xi_1(t)\, x(t - u_1)\, x(t - u_2)\}$$

$$m_1 \{\xi(t)\, x(t - u_1)\, x(t - u_2)\} -$$

$$- \int_{-\infty}^{\infty} h^{*}(u_3)\, m_1 \{x(t - u_1)\, x(t - u_2)\, x(t - u_3)\}\, du_3 =$$

$$m_{\xi x}(u_1, u_2) - \int_{-\infty}^{\infty} h^{*}(u_3)\, m_x(u_1 - u_2,\ u_1 - u_3)\, du_3; \tag{4.87}$$

$$m_x(u_1 - u_2,\ u_1 - u_3) = m_1 \{x(t - u_1)\, x(t - u_2)\, x(t - u_3)\}; \tag{4.88}$$

$$m_x(u_1 - u_2,\ u_1 - u_3,\ u_2 - u_4) =$$

$$= m_1 \{x(t - u_1)\, x(t - u_2)\, x(t - u_3)\, x(t - u_4)\}, \tag{4.89}$$

а $B_{\xi_1}(0)$ совпадает с минимальной среднеквадратической ошибкой, определяемой по формуле (4.13), которая получается при использовании линейных оценок.

Из (4.86) следует, что при использовании в качестве устройства оценки нелинейного фильтра второго порядка среднеквадратическая ошибка оценки зависит уже не только от корреляционных и взаимно корреляционных функций процессов $\xi(t)$ и $\eta(t)$, но и от смешанных моментов *третьего* и *четвертого* порядков.

Для того чтобы найти условие, которому должно удовлетворять ядро $K_2(u_1, u_2)$ нелинейного слагаемого оценки, при котором величина ε_2^2 будет минимальной, исполь-

зуем тот же прием, что и в § 4.2.1. Покажем, что при заданных смешанных моментах процессов до четвертого порядка включительно наилучшая (в смысле принятого критерия минимума среднего квадрата ошибки) нелинейная фильтрация второго порядка процесса $\xi(t)$ из его аддитивной смеси с $\eta(t)$ будет осуществляться при условии, что ядро $K_2(u_1, u_2)$ нелинейного корректирующего члена удовлетворяет следующему двумерному интегральному уравнению:

$$m_{\xi_1 x}(u_1, u_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2^*(u_3, u_4) \times$$
$$\times m_x(u_1 - u_2,\ u_1 - u_3,\ u_2 - u_4)\, du_3\, du_4. \qquad (4.90)$$

Подставим (4.90) в (4.86) и покажем, что ошибка ε_2^2 минимальна при условии $K_2(u, v) = K_2^*(u, v)$. В результате указанной подстановки имеем

$$\varepsilon_2^2 = B_{\xi_1}(0) - 2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2(u_1, u_2) \times$$
$$\times K_2^*(u_3, u_4)\, m_x(u_1 - u_2,\ u_1 - u_3,\ u_2 - u_4)\, du_1\, du_2\, du_3\, du_4 +$$
$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2(u_1, u_2)\, K_2(u_3, u_4) \times$$
$$\times m_x(u_1 - u_2,\ u_1 - u_3,\ u_2 - u_4)\, du_1\, du_2\, du_3\, du_4 =$$
$$= B_{\xi_1}(0) - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2^*(u_1, u_2)\, K_2^*(u_3, u_4) \times$$
$$\times m_x(u_1 - u_2,\ u_1 - u_3,\ u_2 - u_4)\, du_1\, du_2\, du_3\, du_4 +$$
$$+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [K_2(u_1, u_2) - K_2^*(u_1, u_2)]\, [K_2(u_3, u_4) -$$
$$- K_2^*(u_3, u_4)]\, m_x(u_1 - u_2,\ u_1 - u_3,\ u_2 - u_4)\, du_1\, du_2\, du_3\, du_4.$$
$$(4.91)$$

Здесь использовано также очевидное равенство [см. (4.89)]

$$m_x(u_1 - u_2,\ u_1 - u_3,\ u_2 - u_4) = m_x(u_3 - u_4,\ u_3 - u_1,\ u_4 - u_2).$$

Так как только последнее слагаемое в (4.91) содержит неизвестную функцию $K_2(u, v)$ и так как оно неотрица-

тельное, потому что

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[K_2(u_1, u_2)-K_2^*(u_1, u_2)][K_2(u_3, u_4)-$$

$$-K_2^*(u_3, u_4)]\, m_x(u_1-u_2, u_1-u_3, u_2-u_4)\, du_1\, du_2\, du_3\, du_4=$$

$$m_1\left\{\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[K_2(u, v)-K_2^*(u, v)]\, x(t-u)\times\right.\right.$$

$$\left.\left.\times x(t-v)\, du\, dv\right]^2\right\}\geqslant 0,$$

то минимальное значение ε_2^2 будет соответствовать такому нелинейному фильтру второго порядка, для которого указанное слагаемое обращается тождественно в нуль. Это будет иметь место, если ядро $K_2(u, v)$ нелинейного корректирующего члена оценки совпадает с функцией $K_2^*(u, v)$, являющейся решением интегрального уравнения (4.90).

Полагая в (4.91) $K_2(u, v)\equiv K_2^*(u, v)$, найдем минимальное значение среднеквадратической ошибки оценки (4.82):

$$\varepsilon^2_{2\,min}=B_{\xi_1}(0)-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_2^*(u_1, u_2)\, K_2^*(u_3, u_4)\times$$

$$\times m_x(u_1-u_2, u_1-u_3, u_2-u_4)\, du_1\, du_2\, du_3\, du_4 \qquad (4.92)$$

или с учетом (4.90)

$$\xi^2_{2\,min}=\xi^2_{1\,min}-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}K_2^*(u_1, u_2)\, m_{\xi_1 x}(u_1, u_2)\, du_1\, du_2, \qquad (4.93)$$

где

$$\varepsilon^2_{1\,min}=m_1\{\xi_1^2(t)\}=B_{\xi_1}(0)=B_\xi(0)-\int_{-\infty}^{\infty}h^*(u)\, B_{x\xi}(u)\, du$$

— минимальная среднеквадратическая ошибка линейной оценки [см. (4.13)].

Таким образом, использование оптимального нелинейного корректирующего звена в фильтре второго порядка позволяет дополнительно уменьшить среднеквадратическую ошибку при оценке только оптимальной линейной системой

на величину

$$\varepsilon^2_{2\,min}-\varepsilon^2_{1\,min}=\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}K_2^*(u_1,\,u_2)\,m_{\xi_1 x}(u_1,\,u_2)\,du_1\,du_2. \quad (4.94)$$

Рассмотрим задачу об оптимальном фильтре второго порядка в более общем виде, отказавшись от предположения, что линейная часть оценки (4.82) задана. Найдем совместно *две* функции $K_1(u)$ и $K_2(u,\ v)$, минимизирующие средний квадрат ошибки.

Подставляя (4.82) в (4.83), находим развернутое выражение функционала ε_2^2, зависящее от двух неизвестных ядер первого и второго порядка:

$$\varepsilon_2^2=B_\xi(0)-2\int\limits_{-\infty}^{\infty}K_1(u_1)\,B_{\xi x}(u_1)\,du_1-$$

$$-2\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}K_2(u_1,\,u_2)\,m_{\xi x}(u_1,\,u_2)\,du_1\,du_2+$$

$$+\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}K_1(u_1)\,K(u_2)\,B_x(u_2-u_1)\,du_1\,du_2+$$

$$+2\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}K_1(u_1)\,K_2(u_2,\,u_3)\,m_x(u_1-u_2,\,u_1-u_3)\,du_1\,du_2\,du_3+$$

$$+\int\limits_{-\infty}^{\infty}\int\limits_{\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}K_2(u_1,\,u_2)\,K_2(u_3,\,u_4)\times$$

$$\times m_x(u_1-u_2,\ u_1-u_3,\ u_2-u_4)\,du_1\,du_2\,du_3\,du_4. \quad (4.95)$$

Тем же приемом, который применялся выше, можно показать, что минимум среднего квадрата ошибки достигается при использовании такого нелинейного фильтра второго порядка, ядра которого удовлетворяют следующей *системе двух интегральных уравнений*:

$$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}K_2^*(u_2,\,u_3)\,m_x(u_1-u_2,\ u_1-u_3)\,du_2\,du_3+$$

$$+\int\limits_{-\infty}^{\infty}K_1^*(u_2)\,B_x(u_2-u_1)\,du_2=B_{\xi x}(u_1), \quad (4.96)$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2^*(u_3, u_4)\, m_x(u_1-u_2, u_1-u_3, u_2-u_4)\, du_3\, du_4 +$$

$$+\int_{-\infty}^{\infty} K_1^*(u_3)\, m_x(u_1-u_2, u_1-u_3)\, du_3 = m_{\xi x}(u_1, u_2). \quad (4.96')$$

Заметим, что для независимых процессов $\xi(t)$ и $\eta(t)$, имеющих симметричные распределения, $m_x(u_1-u_2, u_1-u_3)=0$ и система уравнений (4.96)—(4.96′) распадается на два уравнения, первое из которых переходит в уравнение (4.12), а второе — в (4.90), так как при этом $m_{\xi x}(u_1, u_2) = m_{\xi_1 x}(u_1, u_2)$. Следовательно, при указанном условии полученное выше решение задачи об оптимальном фильтре второго порядка справедливо и в том случае, когда нет никаких априорных ограничений на линейную часть оценки.

4.4.3. Фильтры произвольного порядка. Дальнейшее улучшение оценки процесса $\xi(t)$ по наблюдаемой реализации $x(t)$ аддитивной смеси этого процесса с другим случайным процессом $\eta(t)$ может быть достигнуто путем использования нелинейных фильтров более высокого порядка. В соответствии с (4.81) оценку, даваемую фильтром n-го порядка, представляем в виде

$$\hat{\xi}(t) = \sum_{i=1}^{n}\int_{-\infty}^{\infty}\dots\int_{-\infty}^{\infty} K_i(u_1, \dots, u_i)\times$$
$$\times x(t-u_1)\dots x(t-u_i)\, du_1\dots du_i. \quad (4.97)$$

При этом возможны две постановки задачи. Можно попытаться отыскать такую последовательность ядер $K_1(u_1), \dots, K_n(u_1, \dots, u_n)$, которая минимизирует средний квадрат ошибки

$$\varepsilon_n^2 = m_1\{[\xi(t)-\hat{\xi}(t)]^2\} = m_1\Big\{\Big[\xi(t)-\sum_{i=1}^{n}\int_{-\infty}^{\infty}\dots$$
$$\dots\int_{-\infty}^{\infty} K_i(u_1, \dots, u_i)\, x(t-u_1)\dots x(t-u_i)\, du_1\dots du_i\Big]^2\Big\}.$$

(4.98)

Условие минимума сводится к системе n интегральных уравнений относительно неизвестных ядер, аналогичной системе (4.96)—(4.96′) для $n = 2$.

Менее общая постановка, приводящая к обозримому результату, аналогична той, которая формулировалась в начале § 4.4.2. Предположим, что $K_1(u)$ совпадает с импульсной переходной функцией оптимальной линейной системы, ядро $K_2(u_1, u_2)$ корректирующего нелинейного элемента второго порядка находим из уравнения (4.90) и ядро $K_3(u_1, u_2, u_3)$ корректирующего элемента третьего порядка определим из условия минимума ε_3^2. Затем добавим нелинейность четвертого порядка и определим $K_4(u_1, u_2, u_3, u_4)$ из условия минимума ε_4^2 и т. д. Найдем рекуррентное уравнение, определяющее $K_n^*(u_1, \ldots, u_n)$, если уже определена последовательность оптимальных ядер до $(n-1)$-го порядка включительно.

Обозначим через $\xi_{n-1}(t)$ ошибку, которая получается, если для оценки $\xi(t)$ используется нелинейный фильтр $(n-1)$-го порядка:

$$\xi_{n-1}(t) = \xi(t) - \sum_{i=1}^{n-1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} K_i(u_1, \ldots, u_i)\, x(t-u_1) \ldots x(t-u_i)\, du_1 \ldots du_i. \tag{4.99}$$

Тогда из (4.98) находим

$$\varepsilon_n^2 = m_1\left\{\left[\xi_{n-1}(t) - \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} K_n(u_1, \ldots, u_n) \times x(t-u_1) \ldots x(t-u_n)\, du_1 \ldots du_n\right]^2\right\} = B_{\xi_{n-1}}(0) - 2\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} K_n(u_1, \ldots, u_n)\, m_{\xi_{n-1}x}(u_1, \ldots, u_n)\, du_1 \ldots du_n + \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} K_n(u_1, \ldots, u_n)\, K_n(u_{n+1}, \ldots, u_{2n}) \times m_x(u_1, \ldots, u_{2n})\, du_1 \ldots du_{2n}, \tag{4.100}$$

где

$$m_{\zeta n-1^x}(u_1, \ldots, u_n) = m_1\{\xi_{n-1}(t)\, x(t-u_1) \ldots x(t-u_n)\}; \tag{4.101}$$

$$m_x(u_1, \ldots, u_{2n}) = m_1\{x(t-u_1) \ldots x(t-u_{2n})\}, \tag{4.102}$$

а $B_{\xi\, n-1}(0)$ совпадает с минимальной среднеквадратической ошибкой, которая может быть достигнута при использовании последовательности нелинейных корректирующих членов до $(n-1)$-го порядка. Для вычисления среднего квадрата ошибки согласно (4.100) требуется уже априорное знание смешанных моментов исходных процессов до $2n$-го порядка включительно.

Следуя применявшемуся выше приему, нетрудно убедиться, что оптимальная по критерию минимума среднего квадрата ошибки нелинейная фильтрация n-го порядка процесса $\xi(t)$ из его аддитивной смеси с $\eta(t)$ будет осуществляться при условии, что ядро $K_n(u_1, \ldots, u_n)$ удовлетворяет следующему многомерному интегральному уравнению:

$$m_{\xi_{n-1}x}(u_1, \ldots, u_n) = \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} K_n^*(u_{n+1}, \ldots, u_{2n})\, m_x(u_1, \ldots, u_{2n})\, du_{n+1} \ldots du_{2n}. \tag{4.103}$$

При $n=2$ уравнение (4.103) переходит в (4.90).

Минимальное значение среднеквадратической ошибки оценки, вырабатываемой нелинейным фильтром n-го порядка, равно

$$\varepsilon^2_{n\,min} = \varepsilon^2_{(n-1)\,min} - \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} K_n^*(u_1, \ldots, u_n) \times m_{\xi_{n-1}x}(u_1, \ldots, u_n)\, du_1 \ldots du_n. \tag{4.104}$$

При $n = 2$ формула (4.104) совпадает с (4.93).

Уравнение (4.103) представляет рекуррентное соотношение, позволяющее принципиально определять характеристику оптимальной корректирующей нелинейности n-го порядка, если уже найдены характеристики оптимальных корректирующих элементов до $(n-1)$-го порядка. Рекуррентной является также формула (4.104), по которой определяется дополнительное уменьшение минимального среднего квадрата ошибки за счет введения нелинейности n-го порядка. Заметим в заключение, что, так же как и в теории оптимальной линейной фильтрации, изложенная выше методика может быть использована и для более широкого класса ситуаций, когда процессы $\xi(t)$ и $\eta(t)$

нестационарны, реализации их суммы наблюдаются на конечном (или полубесконечном) интервале времени, а оценке подвергается значение не процесса $\xi(t)$, а его линейного преобразования.

4.4.4. Фильтрация нормального случайного процесса. До сих пор не делалось никаких специальных предположений о функциях распределения процессов $\xi(t)$ и $\eta(t)$. Предположим теперь, что эти процессы *нормальные*. Тогда нетрудно убедиться, что добавление нелинейных элементов к оптимальному линейному фильтру не может уменьшить величины средней квадратической ошибки.

Как показано выше, оптимальная по критерию минимума среднего квадрата ошибки нелинейная фильтрация 2-го порядка процесса $\xi(t)$ из его аддитивной смеси с $\eta(t)$ осуществляется при условии, что характеристики $K_1(u)$ и $K_2(u, v)$ нелинейного фильтра удовлетворяют системе интегральных уравнений (4.96), (4.96′). Для нормальных процессов $m_x(u_1 - u_2, \ u_1 - u_3) \equiv 0$, $m_{\xi x}(u_1, \ u_2) \equiv 0$ и, следовательно, $K_2^*(u, v) \equiv 0$, т. е. наилучшей в указанном смысле является линейная фильтрация.

Докажем теперь, что добавление нелинейного элемента произвольного порядка также не уменьшает средний квадрат ошибки. Пусть

$$\xi(t) - \hat{\xi}(t) = \xi_1(t) - \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} K_n(u_1, \ \ldots, \ u_n) \times \\ \times x(t-u_1) \ldots x(t-u_n)\, du_1 \ldots du_n, \qquad (4.105)$$

где $\xi_1(t)$ — ошибка, которая получается при использовании оптимальной линейной системы [см. (4.84)]. Повторив рассуждения, которые привели к (4.103), убедимся, что минимум среднего квадрата ошибки (4.105) достигается при условии, что ядро $K_n(u_1, \ldots, u_n)$ удовлетворяет интегральному уравнению

$$m_{\xi_1 x}(u_1, \ \ldots, \ u_n) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} K_n^*(u_{n+1}, \ \ldots, \ u_{2n}) \times \\ \times m_x(u_1, \ \ldots, \ u_{2n})\, du_{n+1} \ldots du_{2n}, \quad n > 1. \qquad (4.106)$$

Как было указано выше, ошибка $\xi_1(t)$ не коррелирована с $x(t)$. Для нормальных случайных процессов $\xi_1(t)$ и $x(t)$ это означает их независимость. Поэтому в рассмат-

риваемом случае

$$m_{\xi_1 x}(u_1, \ldots, u_n) = m_1\{\xi_1(t)\, x(t-u_1) \ldots x(t-u_n)\} = \\ = m_1\{\xi_1(t)\}\, m_1\{x(t-u_1) \ldots x(t-u_n)\},$$

и так как

$$m_1\{\xi_1(t)\} = \int\limits_{-\infty}^{\infty} h^*(u)\, m_1\{x(t-u)\}\, du = 0,$$

то для нормальных процессов с нулевыми средними

$$m_{\xi_1 x}(u_1, \ldots, u_n) \equiv 0. \qquad (4.107)$$

Из (4.106) и (4.107) следует, что

$$K_n^*(u_1, \ldots, u_n) \equiv 0, \quad n > 1. \qquad (4.108)$$

Таким образом, наилучшая фильтрация нормального случайного процесса из аддитивной смеси с нормальным (зависимым, вообще говоря, от него) случайным процессом осуществляется оптимальным линейным фильтром, импульсная переходная функция которого определяется из интегрального уравнения (4.12). Этот результат обобщается и на нестационарные нормальные случайные процессы.

То, что наилучшая фильтрация нормальных процессов линейная, соответствует известному факту совпадения поверхности регрессии для совокупности нормальных случайных величин с плоскостью средней квадратической регрессии (см. приложение V).

4.4.5. Интерпретация нелинейных фильтров. Изложенный в этом разделе метод оптимальной нелинейной фильтрации по критерию минимума среднего квадрата ошибки основан на использовании аппроксимации непрерывного функционала последовательностями вида (4.81). При этом фильтр n-го порядка характеризуется последовательностью ядер $K_n(u_1, \ldots, u_n)$, удовлетворяющих интегральным уравнениям вида (4.103). Решив эти уравнения, можно в принципе по заданной реализации $x(t)$ аддитивной смеси случайных процессов $\xi(t)$ и $\eta(t)$ сформировать согласно (4.97) оценку $\hat{\xi}(t)$ одного из этих процессов, для которой $m_1\{[\hat{\xi}(t) - \xi(t)]^2\}$ минимальна в классе нелинейных фильтров n-го порядка. Практическая реализация фильтра по заданной последовательности ядер K_n связана либо с достаточно сложным вычислительным алгоритмом, либо с какой-либо подходящим образом выбранной интерпретацией

ядер. Одна из таких интерпретаций основана на разложении функции многих переменных в кратные ряды по ортогональным полиномам.

Пусть $\{Q_n(x)\}$ — совокупность ортонормированных полиномов с весовой функцией $\varphi(x)$. Тогда ядро $K_n(u_1, \ldots, u_n)$ можно во многих случаях разложить в n-кратный ряд

$$K_n(u_1, \ldots, u_n) = \varphi(u_1) \ldots \varphi(u_n) \times$$

$$\times \sum_{m_1=0}^{\infty} \ldots \sum_{m_n=0}^{\infty} a_{m_1 \ldots m_n} Q_{m_1}(u_1) \ldots Q_{m_n}(u_n), \; n \geqslant 1, \quad (4.109)$$

коэффициенты которого из условия ортонормированности полиномов равны *)

$$a_{m_1 \ldots m_n} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} K_n(u_1, \ldots, u_n) Q_{m_1}(u_1) \ldots$$

$$\ldots Q_{m_n}(u_n) \varphi(u_1) \ldots \varphi(u_n) \, du_1 \ldots du_n. \quad (4.109')$$

Подставляя (4.109) в (4.81) и обозначая

$$x_k(t) = \int_{-\infty}^{\infty} Q_k(u) \varphi(u) x(t-u) \, du,$$

получаем

$$y_n[x(t)] = K_0 +$$

$$\sum_{k=1}^{n} \sum_{m_1=0}^{\infty} + \ldots \sum_{m_k=0}^{\infty} a_{m_1 \ldots m_k} x_{m_1}(t) \ldots x_{m_k}(t). \quad (4.110)$$

Выражение (4.110) позволяет интерпретировать аппроксимацию нелинейного функционала следующим образом. Имеется набор линейных фильтров с импульсными переходными функциями

$$h_k(u) = \varphi(u) Q_k(u), \quad (4.111)$$

выходные процессы которых перемножаются всеми возможными способами, а затем суммируются с коэффициентами (4.109′), отражающими специфику функционала.

Для физически реализуемых фильтров условие $h_k(u) \equiv 0$ при $u < 0$ выполняется, если используется совокупность

---

*) Для аналогии см. § 2.4.4 в первой книге.

полиномов Лагерра (см. приложение IV в первой книге). В этом случае импульсные переходные функции (4.111) представляют совокупность ортогональных функций Лагерра

$$h_n(u) = \begin{cases} \frac{1}{n!}\frac{d^n}{du^n}(u^n e^{-u}), & u \geqslant 0, \\ 0, & u < 0, \end{cases} \tag{4.112}$$

Приведем в качестве примера приближенное выражение характеристики нелинейного фильтра второго порядка,

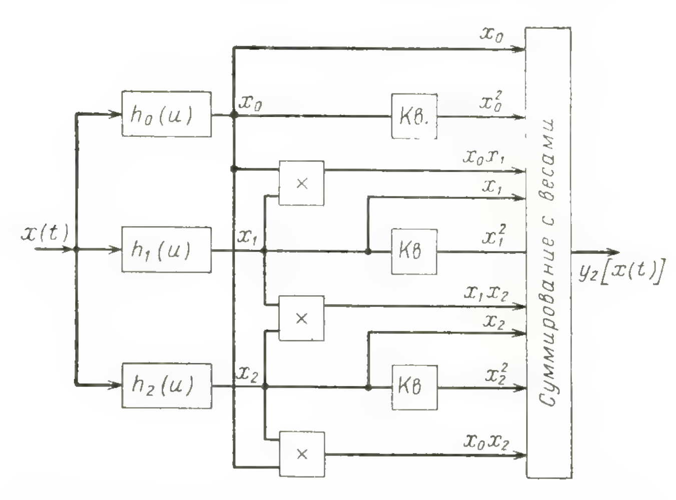


Рис. 24. Блок-схема нелинейного фильтра второго порядка.

если используется набор лагерровских линейных фильтров нулевого, первого и второго порядков. Из (4.110) получим

$$\begin{aligned} y_2[x(t)] \approx K_0 + a_0x_0 + a_1x_1 + a_2x_2 + \\ + a_{00}x_0^2 + (a_{01} + a_{10})\,x_0x_1 + (a_{02} + a_{20})\,x_0x_2 + \\ + a_{11}x_1^2 + (a_{12} + a_{21})\,x_1x_2 + a_{22}x_2^2, \end{aligned} \tag{4.113}$$

где

$$x_0 = x_0(t) = \int_0^\infty e^{-u}x(t-u)\,du; \tag{4.114}$$

$$x_1 = x_1(t) = \int_0^\infty (1-u)\, e^{-u} x(t-u)\, du; \qquad (4.114')$$

$$x_2 = x_2(t) = \int_0^\infty \left(1 - 2u + \frac{u^2}{2}\right) e^{-u} x(t-u)\, du \qquad (4.114'')$$

и

$$a_m = \int_{-\infty}^{\infty} K_1(u)\, h_m(u)\, du, \quad m = 0,\ 1,\ 2, \qquad (4.115)$$

$$a_{mn} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K_2(u, v)\, h_m(u)\, h_n(v)\, du\, dv, \quad m, n = 0,\ 1,\ 2. \qquad (4.115')$$

Блок-схема устройства, реализующего указанный нелинейный фильтр второго порядка, показана на рис. 24.

Задачи

4.1. Пусть $\zeta(t)$ — случайный процесс на выходе линейной системы со случайными параметрами (стационарный в широком смысле), когда на вход ее действует стационарный в широком смысле процесс $\xi(t)$ с энергетическим спектром $F_\xi(\omega)$. Показать, что оптимальная по критерию среднего квадрата ошибки линейная оценка $\zeta(t)$ по реализации процесса $\xi(t)$ осуществляется пропусканием этой реализации через фильтр с передаточной функцией, равной среднему значению $k_c(i\omega)$ передаточной функции указанной выше линейной системы. Доказать, что минимальная среднеквадратическая ошибка при этом равна

$$\varepsilon^2_{min} = \frac{1}{4\pi} \int_{-\infty}^{\infty} F_\xi(\omega)\, [B_c(0,\ \omega) - |k_c(i\omega)|^2]\, d\omega, \qquad (1)$$

где $B_c(\tau,\ \omega)$ — корреляционная функция линейной системы.

4.2. Пусть в постановке задачи § 4.2.3 отбрасывается условие физической осуществимости и оценка процесса (4.36) производится по реализации $x(t)$, наблюдаемой при всех действительных значениях t. Предположим также, что процессы $\xi(t)$ и $\eta(t)$ стационарны в широком смысле, функция $g(t,\ \tau) = g(t - \tau)$ и $G(i\omega)$ — преобразование Фурье функции $g(t)$. Доказать, что передаточная функция оптимального фильтра, вырабатывающего оценку $\zeta(t)$ по реализации суммарного процесса $\xi(t) + \eta(t)$, равна

$$k^*(i\omega) = G(i\omega) \frac{F_{x\xi}(\omega)}{F_x(\omega)}, \qquad (2)$$

а минимальная среднеквадратическая ошибка равна

$$\varepsilon^2_{min} = \frac{1}{2\pi} \int\limits_0^\infty [| G(i\omega)|^2 F_\xi(\omega) - | k^*(i\omega)|^2 F_x(\omega)] \, d\omega =$$

$$= \frac{1}{2\pi} \int\limits_0^\infty | G(i\omega)|^2 \frac{F_\xi(\omega) F_x(\omega) - | F_{x\xi}(\omega)|^2}{F_x(\omega)} \, d\omega. \qquad (3)$$

4.3. Пусть $x(t)$ — реализация случайного процесса, представляющего сумму белого шума с интенсивностью N_0 и независимого от него случайного процесса $\xi(t)$, который получается на выходе интегрирующей RC-цепочки, когда на ее входе действует белый шум с интенсивностью N_s. Показать, что наилучшая оценка по критерию минимума среднего квадрата ошибки экстраполированного на время t_0 значения процесса $\xi(t)$ по реализации $x(t)$ имеет следующий вид:

$$\hat{\xi}(t+t_0) = A_0 \int\limits_0^\infty x(t-u) \, e^{-\alpha \sqrt{1+s^2}\, u} \, du, \quad t_0 > 0, \qquad (4)$$

где

$$A_0 = \frac{\alpha s^2}{1+\sqrt{1+s^2}} e^{-\alpha t_0}; \quad s^2 = \frac{N_s}{N_0}, \quad \alpha = \frac{1}{RC}. \qquad (4')$$

Найти выражение для минимального значения среднего квадрата ошибки экстраполяции

$$\varepsilon^2_{min} = \frac{\alpha N_s}{4} \left[1 - \frac{s^2}{(1+\sqrt{1+s^2})^2} e^{-2\alpha t_0} \right], \quad t_0 > 0. \qquad (5)$$

Получить из (4) и (5) в пределе при $N_0 \to 0$ $(s^2 \to \infty)$ формулы, соответствующие чистой экстраполяции

$$\hat{\xi}(t+t_0) = x(t) \, e^{-\alpha t_0}, \quad t_0 > 0, \qquad (6)$$

$$\varepsilon^2_{min} = \frac{\alpha N_s}{4} (1 - e^{-2\alpha t_0}). \qquad (7)$$

Заметить, что случай фильтрации в момент t получается из приведенных формул при $t_0 = 0$.

4.4. В условиях задачи 4.3 показать, что при $t_0 < 0$ наилучшая оценка интерполированного значения процесса $\xi(t)$ по реализации $x(t)$ имеет следующий вид:

$$\hat{\xi}(t - |t_0|) = \frac{\alpha s^2}{\sqrt{1+s^2}} \int\limits_0^\infty x(t-u) \left[e^{-\alpha \sqrt{1+s^2}\,(u-|t_0|)} - \right.$$

$$\left. - 2\frac{1-\sqrt{1+s^2}}{1+\sqrt{1+s^2}} e^{-\alpha \sqrt{1+s^2}\,(u+|t_0|)} \right] du. \qquad (8)$$

Используя (4.31′), найти для этого случая выражения минимальной среднеквадратической ошибки интерполяции. Убедиться, что при $N_0 \to 0$

$$\hat{\xi}(t - |t_0|) = x(t - |t_0|),$$

т. е. оценка сводится к задержке на время $|t_0|$.

4.5. Показать, что согласованным фильтром для постоянного сигнала $s(t) \equiv a$ является идеальный интегратор с импульсной переходной функцией $h^*(t) = au(t)$, где $u(t)$ — единичный скачок.

4.6. Показать, что для импульсного синусоидального сигнала

$$s(t) = a \sin \omega_0 t, \quad 0 < t < T, \quad \omega_0 T = (2n+1)\pi, \tag{9}$$

импульсная переходная функция согласованного фильтра равна

$$h^*(u) = a \sin \omega_0 u, \; u > 0. \tag{10}$$

ЛИТЕРАТУРА

МОНОГРАФИИ

1. Бендат Дж. Основы теории случайных шумов и ее применения. Пер. с англ., под ред. В. С. Пугачева. Изд-во «Наука», 1965, гл. 4.
2. Вайнштейн Л. А., Зубаков В. Д. Выделение сигналов на фоне случайных помех. Изд-во «Советское радио», 1959, гл. 2, 3.
3. Винер Н. Нелинейные задачи в теории случайных процессов. Пер. с англ., под ред. Ю. Л. Климоновича. Изд-во иностранной литературы, 1961.
4. Давенпорт В. Б., Рут В. Л. Введение в теорию случайных сигналов и шума. Пер. с англ., под ред. Р. Л. Добрушина. Изд-во иностранной литературы, 1960, гл. 11.
5. Дейч Р. Нелинейные преобразования случайных процессов. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1965, гл. 6.
6. Лезин Ю. С. Оптимальные фильтры и накопители импульсных последовательностей. Изд-во «Советское радио», 1963, гл. 1, 2.
7. Ли Р. Оптимальные оценки, определение характеристик и управление. Пер. с англ., под ред. Я. З. Цыпкина. Изд-во «Наука», 1966, гл. 3.
8. Лэннинг Дж. Х., Беттин Р. Г. Случайные процессы в задачах автоматического управления. Пер. с англ., под ред. В. С. Пугачева. Изд-во иностранной литературы, 1958, гл. 7, 8.
9. Миддлтон Д. Введение в статистическую теорию связи, т. II. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1962, гл. 16.
10. Миддлтон Д. Очерки теории связи. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1966, гл. 3.
11. Пугачев В. С. Теория случайных функций и ее применение к задачам автоматического управления, изд. 3. Физматгиз, 1962, гл. 16—18.

12. Солодовников В. В. Статистическая динамика линейных систем автоматического управления. Физматгиз, 1960, гл. 7, 8.
13. Стратонович Р. Л. Условные марковские процессы и их применение к теории оптимального управления. Изд-во Московского университета, 1966.
14. Тихонов В. И. Статистическая радиотехника. Изд-во «Советское радио», 1966, гл. 13.
15. Харкевич В. А. Борьба с помехами. Физматгиз, 1963, § 10, 11.
16. Barrett J. F. Application of the theory of Functionals to Communication Problems. Cambr. Univ., 1955.
17. Wiener N. The interpolation, extrapolation and smoothing of stationary time series. J. Wiley, N. Y., 1949.

СТАТЬИ

18. Боде Г. Шэннон К. Упрощенное изложение линейной минимально-квадратичной теории сглаживания и предсказания. В сб. переводов. «Теория информация и ее приложения». Под ред. А. А. Харкевича, Физматгиз, 1959.
19. Кацнельсон Дж. и Гуд Л. Конструирование нелинейных фильтров и систем управления. В книге Ван-Триса Г. Синтез оптимальных нелинейных систем управления. Пер. с англ., под ред. А. Ю. Ишлинского. Изд-во «Мир», 1964.
20. Колмогоров А. Н. Интерполирование и экстраполирование стационарных случайных последовательностей. «Известия АН СССР», серия математическая, 1941, № 5.
21. Стратонович Р. Л. Применение теории процессов Маркова для оптимальной фильтрации сигналов. «Радиотехника и электроника», 1960, № 11.
22. Ханович И. Г., Яновский Г. Г. Возможные приемы уточнения линейного предсказания стационарных временных рядов. XXII Всесоюзная научная сессия, посвященная дню Радио. Секция теории информации, 1966.
23. Zadeh L. A. On the representation of nonlinear operation. IRE Wescon Conv. Rec., 1957, pt. 2.
24. Zadeh L. A. Progress in Information Theory in USA. IRE Trans. 1961, IT-7, № 2, 1963, IT-9, № 4.

ГЛАВА ПЯТАЯ

ОБНАРУЖЕНИЕ СИГНАЛОВ НА ФОНЕ ПОМЕХ

5.1. ХАРАКТЕРИСТИКА ПРОБЛЕМЫ

Многообразные радиотехнические системы (связи, радиолокации, телеуправления) решают задачи, связанные с передачей информации при помощи электрических сигналов, подверженных искажающему действию различного рода помех. Результатом воздействия помех является частичная или полная потеря полезной информации. В связи с этим возникает ряд проблем, в том числе проблема рационального конструирования (синтеза) системы, состоящая в отыскании оптимальной по некоторому критерию качества процедуры обработки наблюдаемого процесса, являющегося комбинацией сигнала, содержащего в закодированном виде полезную информацию, и маскирующих полезный сигнал помех.

Могут рассматриваться задачи двух видов. Одной из них является задача *обнаружения* сигнала при наличии помех, состоящая в том, чтобы по результатам обработки наблюдаемого процесса, который может быть либо только помехой, либо комбинацией полезного сигнала и помехи, решить, содержится ли полезный сигнал в принятом или нет. Более общей является задача обнаружения одного сигнала (или нескольких) из группы сигналов, наблюдаемых на фоне помех (так называемое *различение* сигналов).

Когда априори известно, что полезный сигнал содержится в наблюдаемом процессе, то возникает задача *выделения* сигнала, скрытого помехой, состоящая в измерении некоторого числа информационных параметров сигнала или описании самого сигнала как функции времени.

Как сигнал, так и искажающие его помехи являются, вообще говоря, случайными процессами, причем наши суждения об этих процессах базируются на выборках конечного размера, представляющих либо конечное множе-

ство чисел (дискретная выборка), либо одну или несколько усеченных реализаций наблюдаемого процесса (непрерывная выборка). Это обстоятельство определяет *статистический подход* к вопросам обработки наблюдаемых процессов с целью извлечения интересующей информации. Точка зрения на прием сигналов как статистическую задачу стала уже традиционной в современной радиотехнике.

Математические основы решения указанных выше задач обнаружения и выделения сигналов на фоне помех содержатся в предыдущих главах, так как формально обнаружение представляет проверку статистических гипотез о характеристиках случайной величины или случайного процесса (главы первая и третья), а выделение — оценку параметров или фильтрацию (главы вторая, третья, четвертая). Также как и проверка гипотез и оценки, обе отмеченные радиотехнические задачи могут рассматриваться с единых позиций теории статистических решений [6, 15].

В этой и следующих главах книги упор будет сделан на детализацию рассмотренных выше статистических методов применительно к радиотехническим системам. Характерной особенностью таких систем является передача информации при помощи модулированных электромагнитных колебаний высокой частоты, вследствие чего процессы на входе приемных устройств, являются *узкополосными* (в том смысле, что ширина полосы спектра процесса много у́же центральной частоты спектра, см. § 4.2.4 в первой книге). Поэтому наряду с процессами, имеющими произвольный спектр, будут отдельно и более подробно рассматриваться узкополосные процессы, спектры которых в основном сосредоточены в полосе частот $\omega_0 \pm \frac{\Delta_c}{2}$, $(\Delta_c \ll \omega_0)$. Как указывалось в § 6.2 первой книги, такие процессы могут быть представлены в виде

$$x(t) = r(t) \cos [\omega_0 t + \vartheta(t)], \tag{5.1}$$

где $r(t)$, $\vartheta(t)$ — медленно меняющиеся огибающая и фаза, спектры которых ограничены практически полосой частот $\left(0, \frac{\Delta_c}{2}\right)$. Поэтому вместо выборки, состоящей из $2N$ мгновенных значений наблюдаемого процесса $(x_1, \ldots, x_{2N})$, можно использовать выборку того же размера, состоящую из N значений огибающей и N значений фазы процесса $(r_1, \vartheta_1, r_2, \vartheta_2, \ldots, r_N, \vartheta_N)$.

Часто бывает желательно до какой-либо специальной обработки принятого сигнала выделить его огибающую или фазу (т. е. *продетектировать*), а затем применять оптимальную процедуру обработки выборки ($r_1, \ldots, r_N$) (амплитудный метод) или ($\vartheta_1, \ldots, \vartheta_N$) (фазовый метод). Поэтому наряду с оптимальными методами обработки поступающих на вход приемного устройства процессов (абсолютно оптимальная *додетекторная* обработка) будут рассмотрены *последетекторные* оптимальные методы обработки. Ясно, что оптимальные амплитудные или фазовые методы не могут быть лучше абсолютно оптимальной процедуры, так как процессы выделения огибающей или фазы, вообще говоря, неизбежно связаны с потерей информации.

Изложение ограничивается случаем, когда помехи — аддитивные и статистически независимые от сигнала шумы, представляющие стационарный нормальный случайный процесс с нулевым средним и энергетическим спектром произвольного вида. Структура полезного сигнала будет оговариваться особо в каждом отдельном случае.

В этой главе рассматриваются задачи обнаружения, а в следующей — выделения сигналов.

5.2. ОПТИМАЛЬНЫЕ АЛГОРИТМЫ ОБНАРУЖЕНИЯ СИГНАЛА В АДДИТИВНОМ НОРМАЛЬНОМ ШУМЕ

5.2.1. Детерминированный сигнал. Начнем с простейшей задачи обнаружения *детерминированного* сигнала *) $s(t)$ в аддитивном нормальном шуме. Задача состоит в проверке простой гипотезы (H_0), что наблюдаемый процесс стационарный, нормальный с нулевым средним, против простой альтернативы (H_1), что этот процесс также нормальный, но со средним значением, изменяющимся по известному точно закону $s(t)$.

Выберем (согласно § 3.4) в качестве наблюдаемых некоррелированные координаты

$$x_k = \sqrt{\lambda_k} \int_{-T}^{T} x(t)\, \varphi_k(t)\, dt, \qquad (5.2)$$

*) Это так называемое *когерентное* обнаружение.

где $x(t)$ — реализация принятого процесса на интервале наблюдения $(-T, T)$, а λ_k и $\varphi_k(t)$ — собственные числа и собственные функции интегрального уравнения

$$\varphi(t) = \lambda \int_{-T}^{T} B(y-t)\,\varphi(y)\,dy, \quad |t| \leqslant T, \tag{5.3}$$

где $B(\tau)$ — известная корреляционная функция шума *).

Ограничиваясь первыми N координатами и учитывая (3.45) и (3.47), нетрудно убедиться, что логарифм отношения правдоподобия для выборки $x_1, \ldots, x_N$ размером N имеет вид

$$\ln l(x_1, \ldots, x_N) = \sum_{i=1}^{N} x_i s_i - \frac{1}{2} \sum_{i=1}^{N} s_i^2, \tag{5.4}$$

$$s_i = \sqrt{\lambda_i} \int_{-T}^{T} s(t)\,\varphi_i(t)\,dt, \quad i = 1, \ldots, N. \tag{5.5}$$

Используя теорию, изложенную в гл. 1, а также учитывая (5.4), нетрудно сформулировать правило выбора решения о наличии или отсутствии сигнала $s(t)$ по наблюдаемым координатам $x_1, \ldots, x_N$ его аддитивной смеси с шумом. Принимается решение (γ_1) о наличии сигнала, если для наблюдаемой выборки

$$\sum_{i=1}^{N} x_i s_i \geqslant \ln c + \frac{1}{2} \sum_{i=1}^{N} s_i^2, \tag{5.6}$$

и принимается решение (γ_0) об отсутствии сигнала, если выполняется неравенство, обратное (5.6).

Таким образом, алгоритм обнаружения сводится к вычислению суммы $\sum_{i=1}^{N} x_i s_i$ и сравнению ее с независящим от наблюдаемой выборки порогом

$$K_N = \ln c + \frac{1}{2} \sum_{i=1}^{N} s_i^2. \tag{5.7}$$

---

*) Результаты, приведенные в гл. 3, позволяют легко обобщить решение рассматриваемой задачи на случай нестационарного нормального шума с нулевым средним и корреляционной функцией $B(t, y)$.

Этот алгоритм является *оптимальным* по любому из рассмотренных в первой главе критериев качества (непоследовательных): байесовскому (включающему как частные случаи максимум апостериорной вероятности и максимум правдоподобия), критерию Неймана - Пирсона и минимаксному. Выбор одного из указанных критериев качества отражается лишь на константе c (см. табл. 1 на стр. 40), т. е. на величине заранее устанавливаемого согласно (5.7) порога K_N.

Как и в общей теории проверки простой гипотезы против простой альтернативы, при обнаружении сигнала на фоне помех использование заранее установленного правила выбора решения связано с возможностью принятия ошибочных решений двух родов. Ошибка первого рода, называемая *ложной тревогой*, состоит в том, что объявляется наличие сигнала, когда его в действительности не было. Ошибка второго рода, называемая *пропуском сигнала*, состоит в том, что объявляется отсутствие сигнала, когда в действительности он содержится в принятом процессе, но маскирован шумом.

Если обозначить через G_1 область N-мерного пространства выборок, удовлетворяющего неравенству (5.6), а через G_0 — остальную часть этого пространства, то условные вероятности ложной тревоги α и пропуска сигнала β запишутся в виде [см. (1.11) и (1.12)]

$$\alpha = P\{\gamma_1 \mid H_0\} = \int \dots \int_{G_1} W_N(x_1, \dots, x_N \mid 0)\, dx_1 \dots dx_N, \tag{5.8}$$

$$\beta = P\{\gamma_0 \mid H_1\} = \int \dots \int_{G_0} W_N(x_1, \dots, x_N \mid s)\, dx_1 \dots dx_N. \tag{5.9}$$

Величину

$$1-\beta = P\{\gamma_1 \mid H_1\} \tag{5.10}$$

называют вероятностью *правильного обнаружения*.

Можно обойти вычисление кратных интегралов в (5.8) и (5.9) и свести его к однократному интегрированию, если воспользоваться общим приемом, указанным в § 1.2.5,

и представить величины α и β в виде

$$\alpha = P\{\ln l(x_1, \ldots, x_N) \geqslant \ln c \mid H_0\} =$$

$$= P\left\{\sum_{i=1}^{N} x_i s_i \geqslant K_N \mid H_0\right\}, \qquad (5.11)$$

$$\beta = P\{\ln l(x_1, \ldots, x_N) < \ln c \mid H_1\} =$$

$$= P\left\{\sum_{i=1}^{N} x_i s_i < K_N \mid H_1\right\}. \qquad (5.12)$$

Случайные величины x_i, как это следует из (5.2), представляют независимые нормальные случайные величины, средние значения которых равны нулю, когда верна гипотеза H_0, и равны s_i, когда верна гипотеза H_1, а дисперсия *всегда* равна

$$M_2\{x_i\} = \lambda_k \int_{-T}^{T}\int_{-T}^{T} B(t-y)\,\varphi_i(t)\,\varphi_i(y)\,dt\,dy =$$

$$= \int_{-T}^{T} \varphi_i^2(t)\,dt = 1.$$

При этом сумма независимых нормальных случайных величин $\sum_{i=1}^{N} x_i s_i$ также будет нормальной случайной величиной с нулевым средним, когда верна гипотеза H_0, со средним значением

$$d_N^2 = \sum_{i=1}^{N} s_i^2, \qquad (5.13)$$

когда верна гипотеза H_1, и с дисперсией, равной d_N^2 всегда.

Тогда из (5.11) и (5.12) находим

$$\alpha = \frac{1}{\sqrt{2\pi d_N^2}} \int_{K_N}^{\infty} e^{-\frac{z^2}{2d_N^2}}\,dz = 1 - F\left(\frac{K_N}{d_N}\right),$$

$$\beta = \frac{1}{\sqrt{2\pi d_N^2}} \int_{-\infty}^{K_N} e^{-\frac{(z-d_N^2)^2}{2d_N^2}}\,dz = F\left(\frac{K_N - d_N^2}{d_N}\right).$$

Учитывая (5.7), можно выражения условных вероятностей ложной тревоги и пропуска сигнала представить в виде

$$\alpha = 1 - F\left(\frac{\ln c}{d_N} + \frac{1}{2} d_N\right), \tag{5.14}$$

$$\beta = F\left(\frac{\ln c}{d_N} - \frac{1}{2} d_N\right). \tag{5.15}$$

Для критерия максимального правдоподобия («идеального наблюдателя») величина c равна единице, и из (5.14), (5.15) следует равенство условных вероятностей:

$$\alpha = \beta = 1 - F\left(\frac{d_N}{2}\right). \tag{5.15'}$$

Отметим, что правило выбора решения, аналогичное (5.6), получается и в том случае, когда в качестве наблюдаемых координат процесса принимаются не независимые величины x_k, определяемые согласно (5.2), а зависимые выборочные значения процесса в N моменты времени, т. е.

$$x_k = x(t_k), \quad k = 1, \ldots, N.$$

Используя матричное представление многомерного нормального распределения (см. (2.57) в первой книге), можно логарифм отношения правдоподобия по *коррелированной* выборке $x_1, \ldots, x_N$ представить в виде

$$\ln l(\mathbf{X}) = \mathbf{s}'\mathbf{M}^{-1}\mathbf{X} - \frac{1}{2}\mathbf{s}'\mathbf{M}^{-1}\mathbf{s}, \tag{5.16}$$

где $\mathbf{X}$ и $\mathbf{s}$ — векторы-столбцы, элементами которых служат $x(t_k)$ и $s(t_k)$ соответственно ($k = 1, \ldots, N$), а $\mathbf{M}^{-1}$ — матрица, обратная корреляционной матрице шума.

Обозначая

$$\mathbf{s}'\mathbf{M}^{-1} = \mathbf{U}, \tag{5.17}$$

можно логарифм отношения правдоподобия представить в виде разности скалярных произведений векторов:

$$\ln l(x_1, \ldots, x_N) = \mathbf{U}\mathbf{X} - \frac{1}{2}\mathbf{U}\mathbf{s} = \sum_{k=1}^{N} u_k x_k - \frac{1}{2}\sum_{k=1}^{N} u_k s_k. \tag{5.18}$$

Правило выбора решения формулируется теперь следующим образом: присутствует сигнал, если для наблюдаемой

выборки

$$\sum_{k=1}^{N} u_k x_k \geqslant \ln c + \frac{1}{2} \sum_{k=1}^{N} u_k s_k, \qquad (5.19)$$

и сигнал отсутствует, если выполняется неравенство, обратное (5.19).

Как и правило (5.6), правило (5.19) предписывает линейную обработку принятой системы выборочных данных — взвешенное суммирование и сравнение результата суммирования с заранее установленным порогом, зависящим от известных априорных характеристик сигнала и шума и от критерия качества. Условные вероятности ложной тревоги и пропуска сигнала определяются по формулам (5.14) и (5.15), в которых вместо величины d_N, подсчитываемой согласно (5.13), подставляется другая, которая находится из соотношения

$$d_N^2 = \sum_{k=1}^{N} u_k s_k. \qquad (5.20)$$

Отличие правила (5.6) от правила (5.19) состоит в том, что в первом случае используются независимые координаты наблюдаемого процесса, получаемые из реализации на заданном интервале интегрированием с весами, которые определяются решением интегрального уравнения (5.3), а во втором — зависимые, непосредственно наблюдаемые выборочные значения, которые затем суммируются с весами, представляющими компоненты произведения вектора-строки сигнала и обратной корреляционной матрицы шума. Наиболее трудоемкой операцией в первом случае является решение интегрального уравнения (5.3), а во втором — обращение корреляционной матрицы **M**. Заметим, что при одинаковом размере выборки N рабочие характеристики обнаружения по алгоритмам (5.6) и (5.19), вообще говоря, отличаются друг от друга.

Наконец, рассмотрим оптимальный алгоритм обнаружения в том случае, когда результаты наблюдений, представленные непрерывной реализацией, не дискретизируются и обработка этих результатов проводится только в аналоговой, а не цифровой форме.

Для этого используем результаты § 3.5, полагая $s_1(t) \equiv s(t)$ и $s_0(t) \equiv 0$. Тогда из (3.81) следует правило выбора решения: сигнал присутствует, если для наблюда-

емой на интервале $(-T,\ T)$ реализации $x\ (t)$

$$\int_{-T}^{T} V(t)\, x(t)\, dt \geqslant \ln c + \frac{1}{2} \int_{-T}^{T} V(t)\, s(t)\, dt = K_T, \quad (5.21)$$

и сигнала нет, если выполняется неравенство, обратное (5.21). Величина c, как и в предыдущих формулах, определяется критерием качества, а функция $V\ (t)$ представляет решение неоднородного линейного интегрального уравнения

$$\int_{-T}^{T} B(t-u)\, V(u)\, du = s(t), \quad |t| \leqslant T. \quad (5.22)$$

Как следует из (3.87) и (3.88), условные вероятности ложной тревоги и пропуска сигнала определяются по формулам (5.14) и (5.15), если в них d_N заменить величиной d_T, которая подсчитывается по формуле [см. (3.86)]

$$d_T = \left(\int_{-T}^{T} V(t)\, s(t)\, dt \right)^{1/2}. \quad (5.23)$$

Нетрудно обнаружить аналогию в формулах (5.6), (5.19), (5.21), а также в формулах (5.13), (5.20) и (5.23).

Заметим, что при обнаружении детерминированного сигнала в аддитивном нормальном белом шуме из (3.79) следует, что функция $V\ (t)$ в формулах (5.21) — (5.23) равна

$$V(t) = \frac{1}{N_0}\, s(t). \quad (5.24)$$

Обозначив через E_s энергию сигнала за интервал наблюдения

$$E_s = \int_{-T}^{T} s^2(t)\, dt, \quad (5.25)$$

получим из (5.21) следующее правило обнаружения сигнала в белом шуме: сигнал присутствует, если

$$\int_{-T}^{T} s(t)\, x(t)\, dt \geqslant \ln c + \frac{E_s}{2N_0}. \quad (5.26)$$

Параметр d_T^2, определяющий условные вероятности ложной тревоги и пропуска сигнала, в этом случае равен

[см. (5.23) — (5.25)]

$$d_T^2 = \frac{E_s}{N_0}, \tag{5.27}$$

т. е. равен отношению энергии сигнала к спектральной плотности шума.

Вводя мощность сигнала $\frac{E_s}{2T}$ и среднюю мощность шума N_0F в эффективной полосе F сигнала, можно формулу (5.27) переписать в виде

$$d_T^2 = \frac{E_s}{N_0} = 2FTs^2, \tag{5.27'}$$

где s^2 — отношение мощности сигнала к средней мощности шума.

Исключая из двух соотношений (5.14) и (5.15) величину $\ln c$, можно записать аналитически связь между вероятностью правильного обнаружения и вероятностью ложной тревоги в виде [ср. (1.77')]

$$x_\alpha - x_{1-\beta} = d, \tag{5.28}$$

где x_α и $x_{1-\beta}$ — процентные точки нормального распределения вероятностей, а параметр d в зависимости от принятой методики обработки определяется по формулам (5.13), (5.20) или (5.23).

На рис. 25 сплошными линиями показано семейство кривых

$$1 - \beta = f(\alpha;\ d), \tag{5.29}$$

называемое рабочей *характеристикой обнаружения*. Штрих-пунктирная линия пересекает кривые в точках, соответствующих критерию максимального правдоподобия («идеального наблюдателя»).

Укажем на важную интерпретацию оптимального алгоритма обнаружения (5.21). Используя результаты § 4.3.1, находим, что оптимальная *линейная* обработка наблюдаемой на интервале $(-T,\ T)$ реализации $x(t)$ для обнаружения детерминированного сигнала $s(t)$ состоит в пропускании этой реализации через согласованный фильтр с импульсной переходной функцией

$$h^*(\tau) = V(T-\tau), \quad |\tau| < T,$$

где $V(t)$ представляет решение интегрального уравнения (5.22). Этот фильтр, как отмечалось в § 4.3.4, может быть

выполнен в виде *коррелометра.* В случае обнаружения в аддитивном белом шуме реализация согласованного фильтра упрощается, так как импульсная переходная функция представляет просто зеркальное отображение сигнала. Соответствующий коррелометр при этом просто вычисляет взаимную временную корреляционную функцию наблюдаемой реализации и детерминированного сигнала.

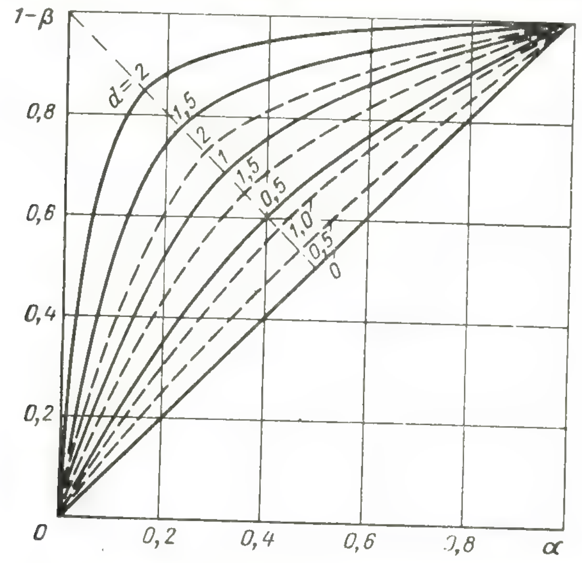


Рис. 25. Рабочие характеристики обнаружения детерминированного (сплошные линии) и квазидетерминированного сигнала (пунктир).

Итак, оптимальное устройство обнаружения детерминированного сигнала на фоне аддитивного нормального шума представляет согласованный фильтр, за которым следует безынерционный пороговый элемент.

5.2.2. Квазидетерминированный сигнал. Рассмотрим теперь более сложную задачу обнаружения *квазидетерминированного* сигнала, представляющего узкополосный процесс с заданными законами амплитудной $a(t)$ и фазовой $\psi_s(t)$ модуляции и со случайной начальной фазой *) φ_0:

$$s(t) = a(t)\cos[\omega_0 t - \psi_s(t) + \varphi_0], \qquad (5.30)$$

*) Это так называемое *некогерентное обнаружение.*

в аддитивном узкополосном нормальном шуме. Задача состоит в проверке простой гипотезы (H_0), что наблюдаемый процесс — стационарный, нормальный с нулевым средним, против *сложной* альтернативы (H_1), что этот процесс также нормальный, но со средним значением $s(t)$, которое представляет одно из континуума реализаций, соответствующих изменению случайной фазы φ_0 в интервале $(-\pi, \pi)$.

Воспользуемся комплексным представлением реализаций узкополосного случайного процесса (см. § 6.2.2 и приложение VII в первой книге)

$$x(t) = \operatorname{Re} z(t) e^{i\omega_0 t}, \tag{5.31}$$

где $z(t)$ — комплексная огибающая узкополосного процесса, связанная с его огибающей $r(t)$ и фазой $\vartheta(t)$ соотношением

$$z(t) = r(t) e^{i\vartheta(t)}, \tag{5.32}$$

причем

$$z(t) = A(t) + iC(t), \tag{5.32'}$$

где $A(t)$ и $C(t)$ — *квадратурные составляющие* узкополосного случайного процесса. Для нормального процесса эти составляющие также нормальны. Комплексная огибающая квазидетерминированного сигнала (5.30), обозначаемая $z_s(t)$, равна

$$z_s(t) = a(t) e^{-i\psi_s(t)}, \tag{5.33}$$

$$s(t) = \operatorname{Re} z_s(t) e^{i\omega_0 t} e^{i\varphi_0}. \tag{5.33'}$$

Теперь реализация $x(t)$ на интервале наблюдения $(-T, T)$ может быть охарактеризована совокупностью некоррелированных координат комплексной огибающей

$$z_k = \sqrt{\lambda_k} \int_{-T}^{T} z(t) \overline{\varphi_k(t)}\, dt. \tag{5.34}$$

Здесь λ_k и $\varphi_k(t)$ — собственные числа и собственные функции интегрального уравнения *)

$$\varphi(t) = \lambda \int_{-T}^{T} B_z(t-u) \varphi(u)\, du, \tag{5.35}$$

*) Нетрудно также убедиться, что функция двух переменных $B_z(t-u)$ симметрична и положительно определенная (см. § 3.3.1), а собственные числа λ_k действительны и положительны.

где

$$B_z(\tau) = \frac{1}{2} m_1 \{ z(t) \overline{z(t+\tau)} \} \tag{5.35'}$$

— корреляционная функция комплексной огибающей шума, действительная и мнимая части которой совпадают с корреляционной и взаимной корреляционной функциями квадратурных составляющих

$$\operatorname{Re} B_z(\tau) = B_A(\tau) = B_C(\tau), \tag{5.36}$$

$$\operatorname{Im} B_z(\tau) = B_{AC}(\tau) = -B_{CA}(\tau). \tag{5.36'}$$

Из (5.34) и (5.35) имеем

$$m_1 \{ z_k \bar{z}_m \} = \begin{cases} 0, & k \neq m, \\ 2, & k = m. \end{cases} \tag{5.37}$$

Используя (5.32), (5.36) и (5.36′), нетрудно показать, что

$$m_1 \{ z_k z_m \} = 0 \tag{5.37'}$$

для всех k и m.

Если x_k и y_k — действительная и мнимая части координаты z_k, то из (5.37), (5.37′) следует (см. § 3.4.5):

$$m_1 \{ x_k y_m \} = 0 \quad \text{для любых} \quad k \text{ и } m, \tag{5.38}$$

$$m_1 \{ x_k^2 \} = m_1 \{ y_k \}^2 = 1, \tag{5.38'}$$

$$m_1 \{ x_k x_m \} = m_1 \{ y_k y_m \} = 0, \quad k \neq m. \tag{5.38''}$$

Кроме того, для шума без сигнала (гипотеза H_0) $m_1 \{ z(t) \} = 0$, и, следовательно,

$$m_1 \{ x_k \mid H_0 \} = m_1 \{ y_k \mid H_0 \} = m_1 \{ z_k \mid H_0 \} = 0, \tag{5.39}$$

а для шума с сигналом (гипотеза H_1) $m_1 \{ z(t) \} = z_s(t) e^{i\varphi_0}$, и поэтому

$$m_1 \{ z_k \mid H_1 \} = s_k e^{i\varphi_0}, \tag{5.39'}$$

$$\begin{aligned} m_1 \{ x_k \mid H_1 \} &= a_k \cos \varphi_0 - b_k \sin \varphi_0, \\ m_1 \{ y_k \mid H_1 \} &= a_k \sin \varphi_0 + b_k \cos \varphi_0, \end{aligned} \tag{5.39''}$$

где

$$s_k = a_k + i b_k = \sqrt{\lambda_k} \int_{-T}^{T} z_s(t) \overline{\varphi_k(t)}\, dt. \tag{5.39'''}$$

Так как компоненты комплексной огибающей нормального узкополосного процесса нормальны, то случайные величины $x_1, y_1, \ldots, x_N, y_N$ представляют совокупность независимых случайных величин. Соответствующее этим $2N$ координатам наблюдаемого процесса выражение логарифма отношения правдоподобия может быть записано в виде [см. (3.71)]

$$\ln l(x_1, y_1, \ldots, x_N, y_N \mid \varphi_0) = \frac{1}{2} \sum_{k=1}^{N} [(x_k^2 + y_k^2) - \\ - (x_k - a_k \cos\varphi_0 + b_k \sin\varphi_0)^2 - (y_k - a_k \sin\varphi_0 - b_k \cos\varphi_0)^2] \tag{5.40}$$

или в комплексной форме

$$\ln l(z_1, \ldots, z_N \mid \varphi_0) = \sum_{k=1}^{N} \left[\operatorname{Re}(z_k \bar{s}_k e^{-i\varphi_0}) - \frac{1}{2} |s_k|^2\right]. \tag{5.41}$$

Переходя к построению правила выбора решения, примем вначале байесовский критерий качества. В этом случае, как было показано в § 1.4.2 [см. (1.110′)], оптимальное правило базируется на сравнении с порогом (зависящим от априорной вероятности наличия сигнала и платы за ошибочные решения) *усредненного* отношения правдоподобия. Так как в рассматриваемой задаче было принято, что случайный параметр φ_0 имеет равномерное распределение, то из (5.41) получаем (см. также § 3.5.3)

$$\Lambda(z_1, \ldots, z_N) = \frac{1}{2\pi} \int_0^{2\pi} l(z_1, \ldots, z_N \mid \varphi_0)\, d\varphi_0 = \\ = \exp\left\{-\frac{1}{2} \sum_{k=1}^{N} |s_k|^2\right\} \frac{1}{2\pi} \int_0^{2\pi} e^{r_N \cos(\varphi_0 - \psi_N)}\, d\varphi_0,$$

где

$$r_N = \left|\sum_{k=1}^{N} z_k \bar{s}_k\right|; \tag{5.42}$$

$$\psi_N = \operatorname{arctg} \frac{\operatorname{Im} \sum_{k=1}^{N} z_k \bar{s}_k}{\operatorname{Re} \sum_{k=1}^{N} z_k \bar{s}_k}. \tag{5.42′}$$

Используя известное интегральное представление функции Бесселя от мнимого аргумента (см., например, первую книгу, стр. 121), получим окончательно

$$\Lambda(z_1, \ldots, z_N) = I_0(r_N) \exp\left\{-\frac{1}{2}\sum_{k=1}^{N} |s_k|^2\right\}. \quad (5.43)$$

Таким образом, если ограничиться первыми N координатами $z_1, \ldots, z_N$ комплексной огибающей наблюдаемого процесса, то байесовское правило выбора решения о наличии или отсутствии квазидетерминированного узкополосного сигнала в аддитивной смеси с шумом может быть сформулировано следующим образом: принимается решение (γ_1) о наличии сигнала, если

$$\ln I_0(r_N) \geqslant \ln c + \frac{1}{2}\sum_{k=1}^{N} |s_k|^2, \quad (5.44)$$

и решение (γ_0) об отсутствии сигнала, если выполняется неравенство, обратное (5.44).

Неравенство (5.44) может быть переписано в виде [ср. с (5.6)]

$$r_N \geqslant c_N, \quad (5.44')$$

где c_N определяется из трансцендентного уравнения

$$\ln I_0(c_N) = K_N = \ln c + \frac{1}{2}\sum_{k=1}^{N} |s_k|^2. \quad (5.45)$$

Итак, оптимальный (байесовский) алгоритм обнаружения в рассматриваемом случае сводится к вычислению величины r_N (т. е. к существенно *нелинейной* обработке наблюдаемых координат) и сравнению результата с независящим от выборки порогом c_N, определяемым согласно (5.45).

Вероятностные характеристики обнаружения зависят прежде всего от распределения случайной величины r_N, определяемой согласно (5.42). Нетрудно найти это распределение. Действительно, случайные величины z_k представляют независимые комплексные нормальные случайные величины с нулевыми средними, когда верна гипотеза H_0, и со средними, равными $s_k e^{i\varphi_0}$, когда верна гипотеза H_1. Дисперсия этих величин для обеих указанных гипотез

равна [см. (2.116) в первой книге]

$$M_2\{|z_k|\} = m_1\{|z_k|^2\} = m_1\{x_k^2\} + m_1\{y_k^2\}$$

или с учетом (5.38′)

$$M_2\{|z_k|\} = 2. \tag{5.46}$$

Отсюда следует, что случайная величина r_N представляет модуль комплексной нормальной случайной величины (или плоского случайного вектора) с независимыми компонентами, дисперсия которой равна $2\sum_{k=1}^{N}|s_k|^2$, а среднее значение равно нулю, когда верна гипотеза H_0, и равно $\sum_{k=1}^{N}|s_k|^2 e^{i\varphi_0}$, когда верна гипотеза H_1.

Распределение модуля такого вектора было подробно рассмотрено в § 3.2.2 первой книги, из которого следует, что случайная величина r_N подчиняется *релеевскому* распределению

$$w_1(r) = \frac{r}{d_N^2} e^{-\frac{r^2}{2d_N^2}}, \quad r > 0, \tag{5.47}$$

когда справедлива гипотеза H_0, и *обобщенному релеевскому* распределению

$$w_1(r) = \frac{r}{d_N^2} e^{-\frac{r^2+d_N^4}{2d_N^2}} I_0(r), \quad r > 0, \tag{5.48}$$

если верна гипотеза H_1.

В формулах (5.47) и (5.48) параметр d_N равен [ср. с (5.13)]

$$d_N^2 = \sum_{k=1}^{N}|s_k|^2. \tag{5.49}$$

Теперь можно записать выражения для вероятностей ложной тревоги и пропуска сигнала. Так как функция $\ln I_0(r)$ монотонно возрастающая, то в соответствии с оптимальным правилом (5.44)

$$\alpha = P\{\ln I_0(r_N) \geqslant K_N | H_0\} = P\{r_N \geqslant c_N | H_0\}, \tag{5.50}$$

$$\beta = P\{\ln I_0(r_N) < K_N | H_1\} = P\{r_N < c_N | H_1\}, \tag{5.51}$$

где c_N — абсцисса точки пересечения кривой $y = \ln I_0(r)$ с прямой $y = K_N$ [см. (5.45)].

Из (5.50) и (5.51) с учетом (5.47) и (5.48) находим

$$\alpha = \int_{c_N}^{\infty} \frac{r}{d_N^2} e^{-\frac{r^2}{2d_N^2}} dr = e^{-\frac{c_N^2}{2d_N^2}}, \qquad (5.52)$$

$$\beta = e^{-\frac{d_N^2}{2}} \int_0^{\frac{c_N}{d_N}} x e^{-\frac{x^2}{2}} I_0(xd_N)\, dx. \qquad (5.53)$$

Последний интеграл подробно табулирован (см. ссылки в третьей главе первой книги).

Наряду с рассмотренной выше процедурой обнаружения по дискретной выборке, найдем аналогичный алгоритм в том случае, когда результаты наблюдений не дискретизируются, а представляются непрерывной реализацией $x(t)$ на интервале $(-T, T)$.

Из (3.78), учитывая (5.35′) и подставляя вместо $s(t)$ величину $z_s(t) e^{i\varphi_0}$, находим выражение для *функционала* отношения правдоподобия, соответствующее наблюдаемой на интервале $(-T, T)$ комплексной огибающей $z(t)$ процесса при заданном φ_0:

$$l[z(t) \mid \varphi_0] = \exp\left\{\operatorname{Re} \int_{-T}^{T} V(t; \varphi_0) \overline{z(t)}\, dt\right\} \times$$

$$\times \exp\left\{-\frac{1}{2} \operatorname{Re} \int_{-T}^{T} V(t; \varphi_0) \overline{z_s(t)} e^{-i\varphi_0} dt\right\}, \qquad (5.54)$$

где $V(t; \varphi_0)$ определяется из неоднородного линейного интегрального уравнения [см. (3.77)]

$$\int_{-T}^{T} B_z(t-u) V(u; \varphi_0)\, du = z_s(t) e^{i\varphi_0}, \qquad |t| < T. \qquad (5.54')$$

Вводя функцию

$$U(t) = V(t; \varphi_0) e^{-i\varphi_0}, \qquad (5.55)$$

можно локализовать параметр φ_0 в выражении функционала отношения правдоподобия. Подставляя (5.55) в (5.54),

получаем для определения $U(t)$ интегральное уравнение

$$\int_{-T}^{T} B_z(t-y)\,U(y)\,dy = z_s(t), \quad |t| \leqslant T, \tag{5.56}$$

из которого видно, что функция $U(t)$ не зависит от φ_0. Заметим, что *комплексное* интегральное уравнение (5.56) эквивалентно системе двух действительных интегральных уравнений относительно действительной $u(t)$ и мнимой $v(t)$ частей функции $U(t)$:

$$\int_{-T}^{T} [B_A(t-y)\,u(y) - B_{AC}(t-y)\,v(y)]\,dy = a(t)\cos\psi_s(t), \tag{5.56'}$$

$$\int_{-T}^{T} [B_A(t-y)\,v(y) + B_{AC}(t-y)\,u(y)]\,dy = a(t)\sin\psi_s(t), \quad |t| \leqslant T. \tag{5.56''}$$

Выражение (5.54) функционала отношения правдоподобия можно с учетом (5.55) переписать в виде

$$\ln[z(t)\,|\,\varphi_0] = \exp\left\{\operatorname{Re} e^{i\varphi_0}\int_{-T}^{T} U(t)\,\overline{z(t)}\,dt\right\} \times$$

$$\times \exp\left\{-\frac{1}{2}\operatorname{Re}\int_{-T}^{T} U(t)\,\overline{z_s(t)}\,dt\right\} =$$

$$= \exp\left\{\left|\int_{-T}^{T} U(t)\,\overline{z(t)}\,dt\right|\cos(\varphi_0-\psi_T)\right\} \times$$

$$\times \exp\left\{-\frac{1}{2}\int_{-T}^{T} U(t)\,\overline{z_s(t)}\,dt\right\}, \tag{5.57}$$

где

$$\psi_T = \operatorname{arctg}\frac{\operatorname{Im}\int_{-T}^{T} U(t)\,\overline{z(t)}\,dt}{\operatorname{Re}\int_{-T}^{T} U(t)\,\overline{z(t)}\,dt}; \tag{5.57'}$$

знак Re при втором сомножителе в (5.57) опущен, так как интеграл действительный и положительный:

$$\int_{-T}^{T} U(t)\overline{z_s(t)}\,dt = \overline{\int_{-T}^{T} \overline{U(t)}\, z_s(t)\,dt} =$$

$$= \overline{\int_{-T}^{T}\int_{-T}^{T} B_z(t-y)\,U(y)\,\overline{U(t)}\,dy\,dt} =$$

$$= m_1\left\{\left|\int_{-T}^{T} \overline{z(t)}\,U(t)\,dt\right|^2\right\} > 0. \qquad (5.58)$$

Усредняя функционал (5.57) по случайной равномерно распределенной (в интервале $|\varphi_0| \leqslant \pi$) фазе, получаем

$$\Lambda[z(t)] = I_0(r_T)\,e^{-\frac{d_T^2}{2}}, \qquad (5.59)$$

где введены обозначения [аналогичные (5.42) и (5.49)]:

$$r_T = \left|\int_{-T}^{T} U(t)\,\overline{z(t)}\,dt\right|; \qquad (5.60)$$

$$d_T^2 = \int_{-T}^{T} U(t)\,\overline{z_s(t)}\,dt. \qquad (5.61)$$

Оптимальное (по байесовскому критерию минимального риска) правило выбора решения может быть теперь сформулировано следующим образом: сигнал присутствует, если для наблюдаемой на интервале $(-T,\ T)$ реализации комплексной огибающей $z(t)$

$$\ln I_0(r_T) \geqslant \ln c + \frac{d_T^2}{2} = K_T \qquad (5.62)$$

или

$$r_T \geqslant c_T, \qquad (5.62')$$

где величина c_T определяется из трансцендентного уравнения

$$\ln I_0(c_T) = K_T. \qquad (5.63)$$

Случайная величина r_T — модуль комплексной нормальной случайной величины $\int_{-T}^{T} U(t)\overline{z(t)}\,dt$, среднее значение которой равно нулю, когда справедлива гипотеза H_0, и равно $d_T^2 e^{i\varphi_0}$, когда справедлива гипотеза H_1. Дисперсия этой величины в соответствии с (5.58) всегда равна [см. (5.35′)]

$$M_2\left\{\left|\int_{-T}^{T} U(t)\overline{z(t)}\,dt\right|\right\}=m_1\left\{\left|\int_{-T}^{T} U(t)\overline{z(t)}\,dt\right|^2\right\}=2d_T^2. \quad (5.64)$$

Таким образом, случайная величина r_T имеет релеевское распределение с параметром d_T^2, когда верна гипотеза H_0, и обобщенное релеевское распределение с параметрами (d_T^1, d_T^2), когда верна гипотеза H_1. Поэтому вероятности ложной тревоги и пропуска сигнала определяются, как и в дискретном случае, также по формулам (5.52), (5.53), с очевидной заменой величин d_N, K_N, c_N на d_T, K_T, c_T, причем последняя определяется из трансцендентного уравнения (5.63).

При обнаружении квазидетерминированного узкополосного сигнала в аддитивном нормальном белом шуме со спектральной плотностью N_0 функция $U(t)$ в формулах (5.60), (5.61) равна

$$U(t)=\frac{1}{N_0}z_s(t). \quad (5.65)$$

Тогда оптимальный алгоритм обнаружения сводится к вычислению величины *)

$$r_T=\frac{1}{N_0}\left|\int_{-T}^{T} z_s(t)\overline{z(t)}\,dt\right| \quad (5.66)$$

и сравнению ее с порогом c_T [см. (5.63)], причем параметр d_T^2 в этом случае равен

$$d_T^2=\frac{1}{N_0}\int_{-T}^{T} |z_s(t)|^2\,dt, \quad (5.66')$$

*) Оптимальное устройство обнаружения может вместо r_T вычислять любую монотонную функцию $f(r_T)$ и сравнивать ее с $f(c_T)$, не изменяя при этом величин вероятностей ошибок. Таким образом, рабочая характеристика обнаружения определяется не отношением сигнал/шум *на выходе* устройства, которое зависит от вида функции $f(r_T)$, а отношением энергии сигнала к спектральной плотности белого шума (или аналогичной величиной (5.61), когда спектр шума неравномерный).

т. е. совпадает с отношением энергии (узкополосного) сигнала к спектральной плотности шума [ср. с (5.27)].

На рис. 25 пунктиром дана рабочая характеристика обнаружения квазидетерминированного узкополосного сигнала. Сравнение с кривыми, изображенными сплошными линиями, при одинаковых значениях параметра d_T дает представление о том, насколько ухудшаются характеристики обнаружения по сравнению со случаем детерминированного сигнала.

В случае обнаружения гармонического сигнала $s(t) = A_0 \cos(\omega_0 t + \varphi_0)$ со случайной фазой из (5.66) получаем

$$r_T = \frac{A_0}{N_0}\left[\left(\int\limits_{-T}^{T} A(t)\,dt\right)^2 + \left(\int\limits_{-T}^{T} C(t)\,dt\right)^2\right]^{\frac{1}{2}}. \qquad (5.67)$$

Учитывая, что квадратурные составляющие $A(t)$ и $C(t)$ наблюдаемой реализации $x(t)$ медленноменяющиеся по сравнению с $\cos\omega_0 t$, (5.67) можно записать в виде *)

$$r_T = \frac{2A_0}{N_0}\left[\left(\int\limits_{-T}^{T} x(t)\cos\omega_0 t\,dt\right)^2 + \right.$$

$$\left. + \left(\int\limits_{-T}^{T} x(t)\sin\omega_0 t\,dt\right)^2\right]^{\frac{1}{2}}. \qquad (5.67')$$

Параметр d_T^2 при этом равен

$$d_T^2 = \frac{2TA_0^2}{N_0}.$$

---

*) Например:

$$\int\limits_{-T}^{T} x(t)\cos\omega_0 t\,dt = \int\limits_{-T}^{T} A(t)\cos^2\omega_0 t\,dt +$$

$$+ \frac{1}{2}\int\limits_{-T}^{T} C(t)\sin 2\omega_0 t\,dt = \frac{1}{2}\int\limits_{-T}^{T} A(t)\,dt + \frac{1}{2}\int\limits_{-T}^{T} [A(t) \times$$

$$\times \cos 2\omega_0 t + C(t)\sin 2\omega_0 t]\,dt \approx \frac{1}{2}\int\limits_{-T}^{T} A(t)\,dt.$$

Заметим, что в соответствии с (3.102) критическая область для фиксированной величины φ_0, в которой отклоняется гипотеза H_0 (т. е. принимается решение о наличии сигнала), определяется по критерию Неймана — Пирсона неравенством

$$\operatorname{Re} \int_{-T}^{T} V(t;\ \varphi_0) \overline{z(t)}\, dt = r_T \cos(\varphi_0 - \psi_T) \geqslant K(\varphi_0), \quad (5.68)$$

где $K(\varphi_0)$ определяется заданной вероятностью α ложной тревоги. Из (5.68) видно, что равномерно наиболее мощного правила в рассматриваемом случае *не существует*. Возможна модификация этого критерия, состоящая в том, что используется усредненное по φ_0 правило (5.68), т. е. байесовское правило (5.62′), в котором c_T находится по формуле [см. (5.52)]

$$c_T = d_T \sqrt{2 \ln \frac{1}{\alpha}}. \quad (5.69)$$

Оптимальный алгоритм обнаружения (5.62′) допускает достаточно простую интерпретацию, если от комплексных величин перейти к действительным. Тогда (5.62′) можно переписать в виде

$$(\operatorname{Re} r_T)^2 + (\operatorname{Im} r_T)^2 \geqslant c_T^2, \quad (5.70)$$

где

$$\operatorname{Re} r_T = \int_{-T}^{T} [u(t) A(t) + v(t) C(t)]\, dt; \quad (5.71)$$

$$\operatorname{Im} r_T = \int_{-T}^{T} [v(t) A(t) - u(t) C(t)]\, dt. \quad (5.71')$$

Пусть $h_c^*(\tau)$ и $h_s^*(\tau)$ — импульсные переходные функции фильтров, согласованных с квадратурными составляющими сигнала $a(t) \cos \psi_s(t)$ и $a(t) \sin \psi_s(t)$. Эти функции определяются решениями системы линейных интегральных уравнений (5.56′) и (5.56″):

$$h_c^*(\tau) = u(T - \tau), \qquad |\tau| \leqslant T, \quad (5.72)$$

$$h_s^*(\tau) = v(T - \tau), \qquad |\tau| \leqslant T. \quad (5.72')$$

Оптимальная *нелинейная* обработка наблюдаемой на интервале $(-T, T)$ реализации $x(t)$ для обнаружения узкополосного квазидетерминированного сигнала в аддитивном

нормальном шуме состоит (рис. 26): 1) в пропускании квадратурных составляющих наблюдаемого процесса через две группы фильтров, согласованных с квадратурными составляющими сигнала, 2) в образовании суммы и разности выходных значений в каждой группе фильтров, 3) в двухполупериодном квадратичном детектировании суммы и разности, 4) в суммировании продетектированных величин.

Если нет фазовой модуляции сигнала, то $h_s^* \equiv 0$, и в каждой из упомянутых групп фильтров остается один фильтр, согласованный с огибающей сигнала $a(t)$.

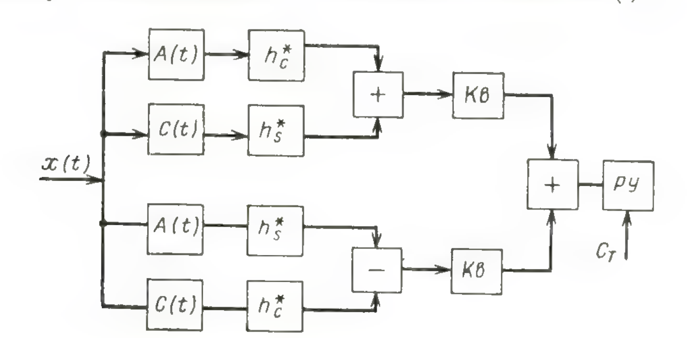


Рис. 26. Блок-схема оптимального устройства обнаружения узкополосного квазидетерминированного сигнала.

Итак, оптимальное устройство обнаружения узкополосного квазидетерминированного сигнала на фоне аддитивного нормального шума состоит из системы фильтров, сумматоров и квадратичных детекторов, за которой следует безынерционный пороговый элемент. Практически используется эквивалентное, но более простое оптимальное устройство, состоящее из одного согласованного с сигналом фильтра и линейного детектора, выделяющего огибающую.

5.2.3. Стохастический сигнал. Предположим, что полезный сигнал, так же как и аддитивный, независимый от сигнала шум, является стационарным нормальным случайным процессом с нулевым средним и отличается от шума только энергетическим спектром (т. е. корреляционной функцией). Пусть известны корреляционные функции $B_c(\tau)$ и $B_ш(\tau)$ сигнала и шума. Задача обнаружения такого стохастического сигнала состоит в проверке гипотезы H_0 о том, что корреляционная функция стационарного нормального случайного процесса равна $B_ш(\tau)$, против альтер-

нативы H_1, что корреляционная функция процесса $B_с(\tau)+B_ш(\tau)$.

Выберем согласно § 3.5.4 в качестве наблюдаемых некоррелированные координаты

$$x_k=\sqrt{\lambda_k}\int_{-T}^{T}x(t)\,\varphi_k(t)\,dt, \qquad (5.73)$$

где $x(t)$ — реализация принятого процесса на интервале $(-T, T)$, а λ_k и $\varphi_k(t)$ — собственные числа и собственные функции (ненормированные) линейного интегрального уравнения [см. (3.120)]

$$\int_{-T}^{T}B_с(t-y)\,\varphi(y)\,dy=(\lambda-1)\int_{-T}^{T}B_ш(t-y)\,\varphi(y)\,dy, \quad |t|\leqslant T, \qquad (5.74)$$

причем нормировка собственных функций производится согласно равенству

$$\lambda_k\int_{-T}^{T}\int_{-T}^{T}B_ш(t-y)\,\varphi_k(t)\,\varphi_m(y)\,dt\,dy=\begin{cases}1, & k=m,\\ 0, & k\neq m.\end{cases} \qquad (5.74')$$

В соответствии с (3.122) логарифм отношения правдоподобия для выборки $x_1, \ldots, x_N$ конечного размера N имеет вид

$$\ln l(x_1, \ldots, x_N)=\frac{1}{2}\sum_{k=1}^{N}\frac{\lambda_k-1}{\lambda_k}x_k^2-\frac{1}{2}\sum_{k=1}^{N}\ln\lambda_k. \qquad (5.75)$$

Правило выбора решения формулируется теперь следующим образом: сигнал присутствует, если

$$\sum_{k=1}^{N}\frac{\lambda_k-1}{\lambda_k}x_k^2\geqslant 2\ln c+\sum_{k=1}^{N}\ln\lambda_k, \qquad (5.76)$$

и отсутствует, если выполняется неравенство, обратное (5.76). Таким образом, алгоритм обнаружения стохастического сигнала сводится к вычислению взвешенной суммы квадратов выборочных данных (некоррелированных координат) и сравнению с заранее устанавливаемым порогом

$$K_N=2\ln c+\sum_{k=1}^{N}\ln\lambda_k. \qquad (5.76')$$

Этот алгоритм *оптимален* по любому из введенных в первой главе (непоследовательных) критериев качества. Выбор того или иного критерия отражается лишь на величине порога (константы c).

Условные вероятности ложной тревоги и пропуска сигнала определяются при помощи распределений (3.126), (3.126'):

$$\alpha = \frac{1}{2\pi} \int_{\ln c}^{\infty} \int_{-\infty}^{\infty} \left[\prod_{k=1}^{N} \left(1 - \frac{iv}{1+v_k}\right)\right]^{-\frac{1}{2}} \times$$

$$\times \exp\left\{-\frac{iv}{2} \ln \prod_{k=1}^{N} \left(1 + \frac{1}{v_k}\right)\right\} e^{ivx}\, dv\, dx, \qquad (5.77)$$

$$\beta = \frac{1}{2\pi} \int_{-\infty}^{\ln c} \int_{-\infty}^{\infty} \left[\prod_{k=1}^{N} \left(1 - \frac{iv}{v_k}\right)\right]^{-\frac{1}{2}} \times$$

$$\times \exp\left\{-\frac{iv}{2} \ln \prod_{k=1}^{N} \left(1 + \frac{1}{v_k}\right)\right\} e^{ivx}\, dv\, dx, \qquad (5.78)$$

где

$$v_k = \frac{1}{\lambda_k - 1}.$$

Интерпретацию алгоритма обнаружения стохастического сигнала в аналоговой форме рассмотрим для случая, когда аддитивный шум *белый*, т. е. $B_{ш}(\tau) = N_0\delta(\tau)$. В этом случае собственные числа λ_k равны

$$\lambda_k = 1 + \frac{1}{N_0\mu_k}, \qquad (5.79)$$

где μ_k — собственные числа линейного однородного интегрального уравнения

$$\psi(t) = \mu \int_{-T}^{T} B_c(t-y)\,\psi(y)\,dy, \qquad (5.80)$$

причем собственные функции $\varphi_k(t)$ исходного уравнения (5.74) связаны с собственными функциями уравнения (5.80) соотношением

$$\psi_k(t) = \sqrt{N_0\lambda_k}\,\varphi_k(t). \qquad (5.80')$$

Наличие аддитивного белого шума исключает сингулярность, и, следовательно, суммы в выражении (5.75) отношения правдоподобия имеют конечные пределы при $N \to \infty$ (см. § 3.5.4). Рассмотрим сумму в левой части (5.76) при $N \to \infty$. Подставляя вместо x_k его выражение из (5.73) и учитывая (5.74), (5.80), (5.80′), получаем

$$\sum_{k=1}^{\infty} \frac{\lambda_k - 1}{\lambda_k} x_k^2 =$$

$$\frac{1}{N_0} \int_{-T}^{T} \int_{-T}^{T} x(u)\, x(v) \left[\sum_{k=1}^{\infty} \frac{\psi_k(u)\, \psi_k(v)}{1 + N_0 \mu_k} \right] du\, dv. \quad (5.81)$$

Функция двух переменных

$$h(u, v) = \sum_{k=1}^{\infty} \frac{\psi_k(u)\, \psi_k(v)}{1 + N_0 \mu_k} \quad (5.82)$$

удовлетворяет интегральному уравнению (см. § 3.5.6)

$$\int_{-T}^{T} B_c(t-u)\, h(u, v)\, du + N_0 h(t, v) = B_c(t-v), \quad (5.83)$$

$$|t| \leqslant T, \qquad |v| \leqslant T.$$

Действительно, подставляя в левую часть (5.83) выражение $h(u, v)$ из (5.82), изменяя порядок суммирования и интегрирования и учитывая (5.80), получаем

$$\int_{-T}^{T} B_c(t-u)\, h(u, v)\, du = \sum_{k=1}^{\infty} \frac{\psi_k(t)\, \psi_k(v)}{\mu_k (1 + N_0 \mu_k)} =$$

$$= \sum_{k=1}^{\infty} \frac{\psi_k(t)\, \psi_k(v)}{\mu_k} - N_0 \sum_{k=1}^{\infty} \frac{\psi_k(t)\, \psi_k(v)}{1 + N_0 \mu_k}.$$

Если учесть, что первая сумма представляет ортогональное разложение корреляционной функции сигнала [см. (3.24)]

$$B_c(t-v) = \sum_{k=1}^{\infty} \frac{\psi_k(t)\, \psi_k(v)}{\mu_k},$$

а вторая сумма в соответствии с (5.82) равна $N_0 h(t, v)$, то справедливость высказанного утверждения становится очевидной.

Используя (5.81) и (5.82), перепишем правило (5.76) выбора решения следующим образом: сигнал присутствует в белом аддитивном шуме, если для наблюдаемой реализации $x(t)$

$$\frac{1}{N_0} \int_{-T}^{T} \int_{-T}^{T} h(u, v) x(u) x(v) \, du \, dv \geqslant K, \tag{5.84}$$

где

$$K = 2 \ln c + \sum_{i=1}^{\infty} \ln \lambda_i \tag{5.84'}$$

и $h(u, v)$ — решение интегрального уравнения (5.83). Бесконечная сумма в (5.84') сходится и выражается через

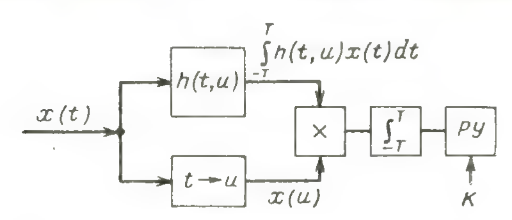


Рис. 27. Блок-схема оптимального устройства обнаружения стохастического сигнала.

резольвенту ядра $B_c(t-u)$ уравнения (5.83) [см. (3.136) в § 3.5.6].

Таким образом, оптимальное устройство обнаружения стохастического (нормального) сигнала на фоне аддитивного нормального белого шума представляет нелинейный фильтр второго порядка (см. § 4.4.2), за которым следует безынерционный пороговый элемент (рис. 27) *).

Вероятности ложной тревоги α и пропуска сигнала β определяются в рассматриваемом случае по формулам (3.141) и (3.141').

В заключение рассмотрим предельный случай, когда сигнал также является нормальным «белым шумом», т. е. случайным процессом с равномерным на всех частотах

*) Ср. также с блок-схемой на рис. 21.

энергетическим спектром, интенсивность которого N_c не равна интенсивности N_0 мешающего белого шума. Тогда $B_c(\tau) = N_c\delta(\tau)$ и из (5.80) и (5.74) следует, что

$$\mu_k = \frac{1}{N_c}, \quad \lambda_k = 1 + \frac{N_c}{N_0}. \qquad (5.85)$$

Правило выбора решения по N некоррелированным выборкам формулируется в этом предельном случае так [см. (5.76)]:

сигнал присутствует, если

$$\frac{N_c/N_0}{1+N_c/N_0} \sum_{k=1}^{N} x_k^2 \geqslant 2\ln c + N\ln\left(1+\frac{N_c}{N_0}\right) \qquad (5.86)$$

или

$$\sum_{k=1}^{N} x_k^2 \geqslant \gamma^2, \qquad (5.87)$$

где

$$\gamma^2 = 2\left(1+\frac{N_0}{N_c}\right)\ln\left[c\left(1+\frac{N_c}{N_0}\right)^{\frac{N}{2}}\right]. \qquad (5.88)$$

При достаточно большом N сумму в левой части неравенства (5.87) можно заменить интегралом *)

$$\sum_{k=1}^{N} x_k^2 \sim \frac{1}{N_0}\int_{-T}^{T} x^2(t)\,dt. \qquad (5.89)$$

Из (5.87) и (5.89) следует, что алгоритм обнаружения сводится к сравнению отношения энергии наблюдаемого процесса $x(t)$ к спектральной плотности шума с некоторой константой γ^2, определяемой согласно (5.88) величиной $\frac{N_c}{N_0}$ и постоянной c (т. е. выбором критерия). Приемное устройство, осуществляющее указанную процедуру, может быть названо *энергетическим приемником*.

Так как x_k — независимые нормальные случайные величины с нулевыми средними, то случайные величины

$$\sum_{k=1}^{N} x_k^2, \quad \frac{N_0}{N_c+N_0}\sum_{k=1}^{N} x_k^2$$

(соответственно, когда сигнала нет и когда сигнал присутствует) имеют χ_N^2-распределение (см. задачу 3.15 в первой

*) Заметим, что величина γ при $N\to\infty$ неограниченна.

книге). Тогда, используя формулу (21) в задаче 3.15, находим условные вероятности ложных тревог и правильного обнаружения:

$$\alpha = 1 - \frac{\Gamma\left(\frac{N}{2},\ \frac{\gamma^2}{2}\right)}{\Gamma\left(\frac{N}{2}\right)}, \tag{5.90}$$

$$\beta = \frac{\Gamma\left(\frac{N}{2},\ \frac{\gamma^2}{2}\frac{N_0}{N_c + N_0}\right)}{\Gamma\left(\frac{N}{2}\right)}, \tag{5.91}$$

где $\Gamma(x, y)$ — неполная гамма-функция [см. (1.41) в первой книге].

Если объем выборки большой ($N \gg 1$), то, используя асимптотическое приближение неполной гамма-функции (см. (1.43) в первой книге), получаем из (5.90) и (5.91)

$$\alpha \sim 1 - F\left[\left(\gamma - \sqrt{N}\right)\sqrt{2}\right], \tag{5.92}$$

$$\beta \sim F\left[\left(\gamma\sqrt{\frac{N_0}{N_c + N_0}} - \sqrt{N}\right)\sqrt{2}\right]. \tag{5.93}$$

Формулы (5.92) и (5.93) связывают между собой (при заданном c) число наблюдений N, отношение мощностей сигнала N_c и шума N_0 на единицу полосы частот и вероятности ложных тревог α и пропуска сигнала β.

5.3. ПОСЛЕДЕТЕКТОРНОЕ ОБНАРУЖЕНИЕ

5.3.1. Амплитудный метод. Переходим к изучению алгоритмов обнаружения при условии, что до какой-либо специальной обработки принятый узкополосный процесс *детектируется*, т. е. выделяется либо его огибающая, либо фаза. При приеме высокочастотные процессы до детектирования усиливаются (например, в усилителе промежуточной частоты), поэтому в этом разделе предполагаем, что ширина полосы пропускания Δ додетекторного усилителя много больше ширины спектра принимаемого сигнала, в отличие от предыдущего раздела, в котором линейная система всегда оказывалась *согласованной* с сигналом.

Рассмотрим сначала, каково оптимальное правило выбора решения о наличии или отсутствии сигнала по реализации *огибающей* наблюдаемого процесса, представляющего

либо нормальный стационарный узкополосный шум с нулевым средним (гипотеза H_0), либо сумму этого шума и детерминированного узкополосного сигнала $s(t) = a(t)\cos[\omega_0 t - \psi_s(t)]$ (гипотеза H_1). Вероятностные характеристики этих процессов приведены в восьмой главе первой книги. Используя указанную там терминологию (см. стр. 427), можно рассматриваемую задачу сформулировать так: проверяется простая гипотеза (H_0), что наблюдаемая огибающая является релеевским процессом, против простой альтернативы (H_1), что она — обобщенный релеевский процесс *).

Следуя общей методике, необходимо было бы в качестве наблюдаемых координат огибающей принять некоррелированные величины

$$r_k = \sqrt{\lambda_k} \int\limits_{T}^{T} r(t)\varphi_k(t)\,dt, \tag{5.94}$$

где $r(t)$ – реализация огибающей на интервале наблюдения $(-T, T)$; λ_k и $\varphi_k(t)$ — собственные числа и собственные функции интегрального уравнения

$$\varphi(t) = \lambda \int\limits_{-T}^{T} B_E(y-t)\varphi(y)\,dy, \qquad |t| \leqslant T, \tag{5.94'}$$

$B_E(\tau)$ — известная корреляционная функция огибающей шума.

Для того чтобы найти функцию распределения случайной величины r_k согласно (5.94), необходимо решить одну из самых сложных задач теории случайных процессов (см. гл. 9 в первой книге), которая состоит в определении распределения процесса на выходе линейной системы, когда распределение процесса на ее входе отлично от нормального (в рассматриваемом случае это распределение релеевское). Случайные величины r_k, $k = 1, 2, \ldots$, не являются ни нормальными, ни релеевскими и из их некоррелированности не следует независимость. Вычисление отношения правдоподобия для выборки $(r_1, \ldots, r_N)$

---

*) Заметим, что распределение огибающей суммы квазидетерминированного процесса и нормального шума также обобщенное релеевское (см. стр. 424 в первой книге). Поэтому выводы этого раздела относятся также к обнаружению квазидетерминированного сигнала [см. (5.30)].

представляет трудно разрешимую задачу. Поэтому отойдем от приведенной *точной* постановки и ценой некоторых допущений упростим задачу *).

Допуская, что энергетический спектр шума равномерный в полосе Δ, представим наблюдаемые некоррелированные координаты огибающей в виде (3.41):

$$r_k = \sqrt{\frac{\Delta}{\pi N_0}} \int_{-T}^{T} r(t) \frac{\sin(\Delta t - \pi k)}{\Delta t - \pi k}\, dt, \qquad (5.95)$$

где N_0 — спектральная плотность шума.

Кроме того, предполагаем, что время наблюдения $T \gg \frac{1}{\Delta}$. Так как при $\Delta \to \infty$ функция $\frac{\Delta}{\pi} \frac{\sin(\Delta t - \pi k)}{\Delta t - \pi k} \to$ $\to \delta\left(t - \frac{\pi k}{\Delta}\right)$, то, учитывая принятое условие $\Delta T \gg 1$, из (5.95) получим

$$r_k = \sqrt{\frac{\pi}{N_0 \Delta}}\, r\left(\frac{\pi k}{\Delta}\right) \qquad (5.96)$$

или

$$r_k = \frac{E_k}{\sigma}, \qquad (5.96')$$

где $\sigma^2 = \frac{N_0 \Delta}{\pi}$ — дисперсия шума;

$E_k = r\left(\frac{\pi k}{\Delta}\right)$ — выборочное значение огибающей при $t - \frac{\pi k}{\Delta}$.

Таким образом, в качестве наблюдаемых координат приняты выборочные значения огибающей через равные интервалы времени $\frac{\pi}{\Delta}$, причем эти значения *приближенно* можно считать некоррелированными.

Так как из некоррелированности значений огибающей нормального случайного процесса следует их статистическая независимость (см. § 8.2 в первой книге), то некоррелированные координаты r_k представляют *независимые* случайные величины.

---

*) Такой упрощенный подход (без достаточно строгих обоснований) был единственным в ранних работах по статистической теории обнаружения сигналов (см., например, [18]).

Ограничиваясь первыми N координатами, нетрудно записать функции правдоподобия выборки $(r_1, \ldots, r_N)$ для двух указанных выше гипотез (см. (8.17) и (8.18) в первой книге):

$$W_N(r_1, \ldots, r_N \mid H_0) = \prod_{k=1}^{N} r_k e^{-\frac{r_k^2}{2}}, \quad r_k > 0, \tag{5.97}$$

$$W_N(r_1, \ldots, r_N \mid H_1) = \prod_{k=1}^{N} r_k e^{-\frac{r_k^2+a_k^2}{2}} I_0(r_k a_k), \quad r_k > 0, \tag{5.97'}$$

где

$$a_k = \frac{a\left(\frac{\pi k}{\Delta}\right)}{\sigma} \tag{5.98}$$

представляет отношение амплитуды сигнала $\left(\text{в момент времени } t = \frac{\pi k}{\Delta}\right)$ к среднеквадратическому значению шума.

Из (5.97) и (5.97′) находим логарифм отношения правдоподобия

$$\ln l\,(r_1, \ldots, r_N) = \sum_{k=1}^{N} \ln I_0(r_k a_k) - \frac{1}{2}\sum_{k=1}^{N} a_k^2. \tag{5.99}$$

Теперь нетрудно сформулировать оптимальное правило выбора решения: если для наблюдаемой выборки $r_1, \ldots, r_N$ координат огибающей

$$\sum_{k=1}^{N} \ln I_0(r_k a_k) \geqslant \ln c + \frac{1}{2}\sum_{k=1}^{N} a_k^2 = K_N, \tag{5.100}$$

то принимается решение (γ_1) о наличии сигнала, а если выполняется неравенство, обратное (5.100), то — решение (γ_0) о его отсутствии.

Константа c в формуле (5.100) определяется принятым критерием качества (байесовским, Неймана — Пирсона, минимаксным) согласно табл. 1 на стр. 40.

Вычисление условных вероятностей ложных тревог и пропусков сигнала в общем виде в этом случае не представляется возможным. Однако поддается исследованию приближенное выражение этих вероятностей при *слабом* сигнале, когда $\max\limits_k a_k \ll 1$.

Разложим $\ln I_0(r_k a_k)$ в степенной ряд, ограничиваясь первым членом, т. е. пренебрегая членами, содержащими степени a_k не ниже четвертой *):

$$\ln I_0(r_k a_k) \approx \frac{1}{4}(r_k a_k)^2.$$

Тогда из (5.100) находим более простое правило: сигнал присутствует, если

$$\sum_{k=1}^{N} a_k^2 r_k^2 \geqslant 4K_N. \tag{5.101}$$

Таким образом, алгоритм обнаружения в рассматриваемом случае сводится к вычислению взвешенной суммы квадратов выборочных значений огибающей и сравнению ее с порогом, зависящим только от выбранного критерия и априорных характеристик сигнала и шума.

Используя (3.6) и (3.7) и возможность их обобщения на случайные процессы (см. § 3.2.3), можно приближенно при $T\Delta \gg 1$ сумму в (5.101) заменить интегралами:

$$\sum_{k=1}^{N} a_k = \frac{1}{\sigma^2}\sum_{k=1}^{N} a^2\left(\frac{\pi k}{\Delta}\right) \approx \frac{E_a}{N_0}, \tag{5.102}$$

где

$$E_a = \int_{-T}^{T} a^2(t)\,dt \tag{5.103}$$

и

$$\frac{1}{4}\sum_{k=1}^{N} r_k^2 a_k^2 = \frac{1}{4\sigma^4}\sum_{k=1}^{N} r^2\left(\frac{\pi k}{\Delta}\right) a^2\left(\frac{\pi k}{\Delta}\right) \approx$$

$$\approx \frac{\Delta}{4\pi\sigma^4}\int_{-T}^{T} r^2(t)\,a^2(t)\,dt. \tag{5.104}$$

Подставляя (5.102) и (5.104) в (5.101), можно правило выбора решения выразить аналитически следующим неравенством:

$$\frac{1}{4N_0\sigma^2}\int_{-T}^{T} r^2(t)\,a^2(t)\,dt \geqslant \ln c + \frac{E_a}{2N_0}, \tag{5.105}$$

---

*) Разложение в ряд и сделанное приближение должны интерпретироваться в вероятностном смысле.

из которого следует, что основными элементами оптимального приемного устройства при амплитудном методе обнаружения слабого сигнала являются квадратичный детектор и коррелометр, вычисляющий взаимную корреляцию выхода $r^2(t)$ детектора и квадрата огибающей $a^2(t)$ детерминированного сигнала.

Для определения вероятностей ложных тревог и пропуска сигнала при использовании алгоритма (5.101) или (5.105) необходимо в общем случае решить достаточно сложную задачу о распределении взвешенной суммы квадратов независимых случайных величин, распределенных по обобщенному закону Релея, в первом случае, или задачу о распределении интеграла от квадрата обобщенного релеевского процесса, во втором.

Если размер выборки большой ($N \gg 1$), то распределение суммы в левой части (5.100) независимых случайных величин можно в силу центральной предельной теоремы при принятых ограничениях относительно a_k (см. § 3.4 в первой книге) считать нормальным со средним и дисперсией, равными *) (при $\max\limits_k a_k \ll 1$)

$$m_1\left\{\sum_{k=1}^{N} \ln I_0(a_k r_k) \,|\, H_0\right\} \approx 2\sum_{k=1}^{N} a_k^2\left(1-\frac{a_k^2}{2}\right), \quad (5.106)$$

$$M_2\left\{\sum_{k=1}^{N} \ln I_0(a_k r_k) \,|\, H_0\right\} \approx 4\sum_{k=1}^{N} a_k^4, \quad (5.106')$$

когда сигнала нет, и

$$m_1\left\{\sum_{k=1}^{N} \ln I_0(a_k r_k) \,|\, H_1\right\} \approx 2\sum_{k=1}^{N} a_k^2\left(1+\frac{a_k^2}{2}\right), \quad (5.106'')$$

$$M_2\left\{\sum_{k=1}^{N} \ln I_0(a_k r_k) \,|\, H_1\right\} \approx 4\sum_{k=1}^{N} a_k^4, \quad (5.106''')$$

когда сигнал присутствует.

Из (5.106) — (5.106''') следует, что при $\max\limits_k a_k \ll 1$

*) О вычислении моментов для квадратов релеевской и обобщенной релеевской случайных величин см. в гл. 2 и 3 первой книги.

и $N \gg 1$ логарифм отношения правдоподобия (5.99) распределен (приближенно) по нормальному закону с параметрами:

$$m_1\{\ln l(r_1, \ldots, r_N) \mid H_0\} = -\frac{1}{4}\sum_{k=1}^{N} a_k^4, \qquad (5.107)$$

$$m_1\{\ln l(r_1, \ldots, r_N) \mid H_1\} = \frac{1}{4}\sum_{k=1}^{N} a_k^4, \qquad (5.107')$$

$$M_2\{\ln l(r_1, \ldots, r_N) \mid H_0\} =$$

$$= M_2\{\ln l(r_1, \ldots, r_N) \mid H_1\} = \frac{1}{4}\sum_{k=1}^{N} a_k^4. \qquad (5.107'')$$

Обозначив

$$s_N^2 = \frac{1}{4}\sum_{k=1}^{N} a_k^4, \qquad (5.108)$$

представим вероятности ложных тревог и пропусков сигнала следующими формулами:

$$\alpha = 1 - F\left(\frac{\ln c}{s_N} + s_N\right), \qquad (5.109)$$

$$\beta = F\left(\frac{\ln c}{s_N} - s_N\right). \qquad (5.109')$$

Заметим, что для абсолютно оптимального метода обнаружения вероятности ложных тревог и пропуска сигнала [см. (5.14) и (5.15)] определяются параметром d_N, который равен корню квадратному из суммы квадратов нормированных значений сигнала, в то время как указанные величины в формулах (5.109) и (5.109′) определяются параметром s_N, который равен корню квадратному из суммы четвертых степеней нормированных значений огибающей сигнала. Отсюда следует, что при слабом сигнале ухудшение рабочей характеристики обнаружения при оптимальной последетекторной обработке по сравнению с оптимальной додетекторной обработкой связано с тем, что в первом случае она определяется *квадратом* отношения сигнал/шум, а во втором — *первой степенью* этого отношения.

Аналогичным образом можно исследовать асимптотические свойства рабочей характеристики обнаружения, исхо-

дя из алгоритма (5.105) обработки реализации при $T \to \infty$ и учитывая, что при этом распределение интеграла в (5.105) приближается к нормальному (см. § 9.4.2 в первой книге).

Заметим, что для *сильного* сигнала, т. е. при $\min\limits_k a_k \gg 1$, используя асимптотическое разложение бесселевой функции (см., например, стр. 123 в первой книге), неравенство (5.100) можно представить в виде

$$\sum_{k=1}^{N} a_k r_k \geqslant K_N \tag{5.110}$$

или приближенно при $T\Delta \gg 1$ [см. (5.105)] в виде

$$\frac{1}{N_0} \int_{-T}^{T} r(t)\, a(t)\, dt \geqslant \ln c + \frac{E_a}{2N_0}. \tag{5.111}$$

В этом случае оптимальное приемное устройство содержит линейный детектор и коррелометр, определяющий взаимную корреляцию между выходом детектора и огибающей детерминированного сигнала.

При $N \gg 1$ аналогично предыдущему находим, что логарифм отношения правдоподобия при $\min\limits_k a_k \gg 1$ имеет нормальное распределение с параметрами:

$$m_1 \{\ln l(r_1, \ldots, r_N) \,|\, H_0\} \sim -\frac{1}{2} \sum_{k=1}^{N} a_k^2. \tag{5.112}$$

$$m_1 \{\ln l(r_1, \ldots, r_N) \,|\, H_1\} \sim \frac{1}{2} \sum_{k=1}^{N} a_k^2, \tag{5.112'}$$

$$M_2 \{\ln l(r_1, \ldots, r_N) \,|\, H_0\} \sim 2 \sum_{k=1}^{N} a_k^2, \tag{5.112''}$$

$$M_2 \{\ln l(r_1, \ldots, r_N) \,|\, H_1\} \sim \sum_{k=1}^{N} a_k^2. \tag{5.112'''}$$

Вероятности ложных тревог и пропусков сигнала равны соответственно

$$\alpha = 1 - F\left(\frac{\ln c}{d_N \sqrt{2}} + \frac{d_N}{2\sqrt{2}}\right), \tag{5.113}$$

$$\beta = F\left(\frac{\ln c}{d_N} - \frac{d_N}{2}\right), \tag{5.114}$$

где

$$d_N^2 = \sum_{k=1}^{N} a_k^2. \qquad (5.115)$$

Заметим, что в рассматриваемом случае сильного сигнала ($d_N \gg 1$) вероятность пропуска сигнала, подсчитанная согласно (5.114), не отличается от этой величины при оптимальной додетекторной обработке [см. (5.15)]. При определении вероятности ложной тревоги второе слагаемое (основное) в аргументе интеграла Лапласа в $\sqrt{2}$ раз меньше аналогичного слагаемого в формуле (5.14).

5.3.2. Фазовый метод. Перейдем к определению оптимального правила выбора решения о наличии или отсутствии детерминированного сигнала в аддитивном узкополосном нормальном шуме по реализации *фазы* $\vartheta(t)$ наблюдаемого процесса. Вероятностные характеристики фазы, соответствующие двум гипотезам H_0 (фаза шума) и H_1 (фаза аддитивной смеси сигнала и шума), приведены в восьмой главе первой книги.

Проводя рассуждения, аналогичные тем, которые предшествовали формуле (5.96), введем некоррелированные координаты фазы

$$\vartheta_k = \vartheta\left(\frac{\pi k}{\Delta}\right). \qquad (5.116)$$

Так как из некоррелированности значений фазы нормального случайного процесса следует их независимость (см. § 8.4.2 в первой книге), то координаты ϑ_k независимы. Ограничиваясь первыми N координатами, запишем функции правдоподобия выборки $\vartheta_1, \ldots, \vartheta_N$ для двух указанных гипотез (см. (8.59) в первой книге):

$$w_N(\vartheta_1, \ldots, \vartheta_N | H_0) = \left(\frac{1}{2\pi}\right)^N, \quad |\vartheta_i| \leqslant \pi, \qquad (5.117)$$

$$w_N(\vartheta_1, \ldots, \vartheta_N | H_1) = \prod_{k=1}^{N} \left\{ \frac{1}{2\pi} e^{-\frac{a_k^2}{2}} + \right.$$

$$+ \frac{a_k}{\sqrt{2\pi}} \cos(\vartheta_k - \psi_k) F[a_k \cos(\vartheta_k - \psi_k)] \times$$

$$\left. \times e^{-\frac{a_k^2}{2} \sin^2(\vartheta_k - \psi_k)} \right\}, \quad |\vartheta_i - \psi_i| \leqslant \pi, \; i = 1, \ldots, N, \qquad (5.118)$$

где a_k определяется по формуле (5.98), а

$$\psi_k = \psi_s\left(\frac{\pi k}{\Delta}\right). \tag{5.119}$$

Из (5.117) и (5.118) находим логарифм отношения правдоподобия

$$\ln(\vartheta_1, \ldots, \vartheta_N) = \sum_{k=1}^{N} \ln\{e^{-\frac{a_k^2}{2}} + \sqrt{2\pi}\, a_k \cos(\vartheta_k - \psi_k) \times$$

$$\times F[a_k \cos(\vartheta_k - \psi_k)]\, e^{-\frac{a_k^2}{2}\sin^2(\vartheta_k - \psi_k)}\}, \tag{5.120}$$

$$|\vartheta_i - \psi_i| \leqslant \pi, \quad i = 1, \ldots, N.$$

Правило выбора решения может быть сформулировано так: сигнал присутствует, если

$$\sum_{k=1}^{N} \ln\{e^{-\frac{a_k^2}{2}} + \sqrt{2\pi}\, a_k \cos(\vartheta_k - \psi_k) \times$$

$$\times F[a_k \cos(\vartheta_k - \psi_k)]\, e^{-\frac{a_k^2}{2}\sin^2(\vartheta_k - \psi_k)}\} \geqslant \ln c, \tag{5.121}$$

$$|\vartheta_i - \psi_i| \leqslant \pi, \quad i = 1, \ldots, N.$$

Для *слабого* сигнала ($\max\limits_k a_k \ll 1$) правило (5.121) значительно упрощается. Используя формулу (8.61) первой книги, находим в первом приближении

$$\sqrt{\frac{\pi}{2}} \sum_{k=1}^{N} a_k \cos(\vartheta_k - \psi_k) \geqslant \ln c + \frac{\pi}{8} \sum_{k=1}^{N} a_k^2 = K_N, \tag{5.122}$$

$$|\vartheta_i - \psi_i| \leqslant \pi, \quad i = 1, \ldots, N.$$

Как и в предыдущих случаях, при $\Delta T \gg 1$ сумму в левой части (5.122) можно приближенно заменить интегралом

$$\frac{\Delta}{\pi\sigma} \int_{-T}^{T} a(t) \cos[\vartheta(t) - \psi_s(t)]\, dt \geqslant \sqrt{\frac{2}{\pi}} \ln c + \sqrt{\frac{\pi}{32}} \frac{E_a}{2N_0}. \tag{5.122'}$$

Таким образом, основной операцией в оптимальной процедуре обнаружения слабого детерминированного сигнала фазовым методом является вычисление взаимной корреля-

ции между косинусом разности фаз принятого и детерминированного сигналов и огибающей детерминированного сигнала *).

Как и в амплитудном методе, определение вероятностей ложных тревог и пропусков сигнала при использовании алгоритма обработки (5.122) сопряжено с трудностью вычисления функции распределения суммы случайных величин в левой части (5.122). Однако если размер выборки большой ($N \gg 1$), то можно воспользоваться асимптотической нормальностью указанной суммы, ограничившись при этом вычислением среднего и дисперсии для двух гипотез. Получим при $\max_k a_k \ll 1$ (см. § 8.5.1 в первой книге):

$$m_1\left\{\sqrt{\frac{\pi}{2}}\sum_{k=1}^{N} a_k\cos(\vartheta_k-\psi_k)\,|\,H_0\right\}=0, \qquad (5.123)$$

$$m_1\left\{\sqrt{\frac{\pi}{2}}\sum_{k=1}^{N} a_k\cos(\vartheta_k-\psi_k)\,|\,H_1\right\}=\frac{\pi}{4}\sum_{k=1}^{N} a_k^2, \qquad (5.123')$$

$$M_2\left\{\sqrt{\frac{\pi}{2}}\sum_{k=1}^{N} a_k\cos(\vartheta_k-\psi_k)\,|\,H_0\right\}=$$

$$=M_2\left\{\sqrt{\frac{\pi}{2}}\sum_{k=1}^{N} a_k\cos(\vartheta_k-\psi_k)\,|\,H_1\right\}=\frac{\pi}{4}\sum_{k=1}^{N} a_k^2. \qquad (5.123'')$$

Используя обозначение (5.115), запишем выражения для вероятностей ложных тревог и пропусков сигнала в виде

$$\alpha=1-F\left(\frac{2\ln c}{\sqrt{\pi}\,d_N}+\frac{\sqrt{\pi}}{4}d_N\right), \qquad (5.124)$$

$$\beta=F\left(\frac{2\ln c}{\sqrt{\pi}\,d_N}-\frac{\sqrt{\pi}}{4}d_N\right). \qquad (5.124')$$

В отличие от оптимального амплитудного метода [см. (5.109) и (5.109')], при оптимальном фазовом методе, как и при оптимальном додетекторном методе обнаружения, вероятности ложных тревог и пропусков сигнала определяют-

*) Как показано в [23], оптимальное устройство при обнаружении фазовым методом представляет идеальный ограничитель, за которым следует согласованный фильтр.

ся параметром d_N, т. е. рабочая характеристика обнаружения слабого сигнала при фазовом методе оказывается лучше *), чем при амплитудном. Нетрудно обнаружить полную аналогию (5.124), (5.124') с (5.14), (5.15), причем первые из указанных формул получаются из вторых заменой d_N на $\sqrt{\frac{\pi}{4}}\, d_N$ (т. е. $\approx 0{,}9\, d_N$).

Если воспользоваться асимптотической формулой (8.63) в первой книге, то для *сильного* сигнала ($\min\limits_k a_k \gg 1$) правило выбора решения [см. (5.121)] может быть записано в виде: сигнал присутствует, если

$$\frac{1}{2}\sum_{k=1}^{N} a_k^2 (\vartheta_k - \psi_k)^2 < \sum_{k=1}^{N} \ln\left(\sqrt{2\pi}\, a_k c^{-\frac{1}{N}}\right) \quad (5.125)$$

или приближенно при $\Delta T \gg 1$

$$\frac{1}{2N_0}\int_{-T}^{T} a^2(t)\,[\vartheta(t) - \psi_s(t)]^2\, dt \leqslant$$

$$\leqslant \frac{\Delta}{\pi}\int_{-T}^{T} \ln\left[\sqrt{2\pi}\,\frac{a(t)}{\sigma}\right] dt - \ln c. \quad (5.125')$$

Таким образом, основными элементами оптимального приемного устройства при фазовом методе обнаружения сильного детерминированного сигнала является фазовый детектор, выделяющий квадрат разности фаз принятого и детерминированного сигналов, и коррелометр, определяющий взаимную корреляцию выхода фазового детектора и квадрата огибающей сигнала.

Рассмотрим также оптимальный фазовый метод обнаружения квазидетерминированного сигнала [см. (5.30)]. В этом случае для слабого сигнала ($\max\limits_k a_k \ll 1$) отноше-

---

*) Этого следовало ожидать, так как функция распределения фазы (5.118) информативнее функции распределения огибающей (5.97'). В первом случае параметры распределения содержат информацию о значениях огибающей и фазы сигнала, а во втором — только о значениях огибающей сигнала.

ние правдоподобия при фиксированном значении фазы φ_0 имеет вид [см. (5.122)]

$$l(\vartheta_1, \ldots, \vartheta_N \mid \varphi_0) =$$

$$= \exp\left\{\sqrt{\frac{\pi}{2}} \sum_{k=1}^{N} a_i \cos(\vartheta_k - \psi_k + \varphi_0) - \frac{\pi}{8} \sum_{k=1}^{N} a_k^2\right\} \qquad (5.126)$$

или

$$l(\vartheta_1, \ldots, \vartheta_N \mid \varphi_0) = \exp\left\{\sqrt{\frac{\pi}{2}}\, r_N \cos(\varphi_0 - \Phi_N)\right\} \times$$

$$\times \exp\left\{-\frac{\pi}{8} \sum_{k=1}^{N} a_k^2\right\}, \qquad (5.126')$$

где

$$r_N = \left| \sum_{k=1}^{N} a_k e^{i(\vartheta_k - \psi_k)} \right|; \qquad (5.127)$$

$$\Phi_N = \operatorname{arctg} \frac{\operatorname{Im} \sum_{k=1}^{N} a_k e^{i(\vartheta_k - \psi_k)}}{\operatorname{Re} \sum_{k=1}^{N} a_k e^{i(\vartheta_k - \psi_k)}}. \qquad (5.127')$$

Усредняя отношение правдоподобия по равномерно распределенной случайной фазе φ_0, получаем

$$\Lambda(\vartheta_1, \ldots, \vartheta_N) = \exp\left\{-\frac{\pi}{8} \sum_{k=1}^{N} a_k^2\right\} \times$$

$$\times \frac{1}{2\pi} \int_0^{2\pi} \exp\left\{\sqrt{\frac{\pi}{2}}\, r_N \cos(\varphi_0 - \Phi_N)\right\} d\varphi_0 =$$

$$= I_0\left(\sqrt{\frac{\pi}{2}}\, r_N\right) \exp\left\{-\frac{\pi}{8} \sum_{k=1}^{N} a_k^2\right\}. \qquad (5.128)$$

Из (5.128) следует байесовское правило выбора решения: квазидетерминированный сигнал присутствует, если

$$\ln I_0\left(\sqrt{\frac{\pi}{2}}\, r_N\right) \geqslant \ln c + \frac{\pi}{8} \sum_{k=1}^{N} a_k^2 = K_N$$

или если

$$r_N \geqslant c_N, \qquad (5.128')$$

где c_N — величина, определяемая из трансцендентного уравнения

$$\ln I_0\left(\sqrt{\frac{\pi}{2}}\, c_N\right) = K_N. \qquad (5.128'')$$

Вычисление функции распределения статистики r_N [см. (5.127)] можно выполнить в замкнутом виде для гипотезы H_0, т. е. когда сигнала нет. В этом случае, используя несколько видоизмененный метод, указанный в [5, гл. 4, § 4] для задачи о случайных блужданиях, получим следующее выражение одномерной функции распределения r_N, когда ϑ_k — независимые, равномерно распределенные на интервале $(0,2\pi)$ случайные величины:

$$w_1(r) = r\int_0^\infty sJ_0(rs)\prod_{k=1}^{N} J_0(a_k s)\,ds.$$

Отсюда с учетом (5.128′) следует, что вероятность ложной тревоги равна

$$\alpha = \int_{c_N}^\infty w_1(r)\,dr = 1 - c_N\int_0^\infty J_1(c_N s)\prod_{k=1}^{N} J_0(a_k s)\,ds. \qquad (5.129)$$

Для синусоидального сигнала постоянной амплитуды

$$\alpha = 1 - k\int_0^\infty J_1(kx)\,J_0^N(x)\,dx, \qquad (5.129')$$

где $k = \frac{c_N}{A_0}$. При использовании критерия Неймана — Пирсона порог k определяется из (5.129′) по заданной вероятности α ложных тревог и не зависит от мощности сигнала.

Если размер выборки велик ($N \gg 1$), то распределение случайной величины r_N асимптотически приближается к релеевскому, когда верна гипотеза H_0, и к обобщенному релеевскому, когда верна гипотеза H_1 (см. [5, стр. 187—188]). Параметры этих распределений равны соответственно [см. (5.123) — (5.123″)] $\frac{1}{2}d_N^2$ и $\left(\sqrt{\frac{\pi}{8}}\, d_N^2,\ \frac{1}{2}d_N^2\right)$, где d_N^2 опре-

деляется по формуле (5.115). Тогда вероятности ложных тревог и пропусков сигнала равны [ср. (5.52) и (5.53)]

$$\alpha = e^{-\frac{c_N^2}{d_N^2}}, \tag{5.130}$$

$$\beta = e^{-\frac{\pi}{8}d_N^2} \int_0^{\frac{c_N\sqrt{2}}{d_N}} x e^{-\frac{x^2}{x}} I_0\left(\sqrt{\frac{\pi}{4}}\, x d_N\right) dx. \tag{5.130'}$$

5.3.3. Стохастический сигнал. Рассмотрим теперь оптимальный амплитудный метод обнаружения стохастического сигнала, сохранив предположения, сделанные в начале раздела 5.3. Функции правдоподобия в этом случае имеют вид [см. (5.97)]

$$W_N(r_1, \ldots, r_N \mid H_0) = \prod_{k=1}^{N} r_k e^{-r_k^2/2}, \quad r_k \geqslant 0,$$

$$W_N(r_1, \ldots, r_N \mid H_1) = \prod_{k=1}^{N} \frac{r_k}{1 + \frac{\sigma_c^2}{\sigma^2}} e^{-\frac{r_k^2}{2\left(1+\frac{\sigma_c^2}{\sigma^2}\right)}},$$

$$r_k \geqslant 0, \quad k = 1, \ldots, N,$$

где σ_c^2 — дисперсия сигнала.

Отношение правдоподобия равно

$$l(r_1, \ldots, r_N) = \left(1 + \frac{\sigma_c^2}{\sigma^2}\right)^{-N} \times$$

$$\times \exp\left\{\frac{\sigma_c^2/\sigma^2}{2(1+\sigma_c^2/\sigma^2)} \sum_{k=1}^{N} r_k^2\right\}, \tag{5.131}$$

$$r_k \geqslant 0, \quad k = 1, \ldots, N.$$

Правило решения формулируется так: сигнал присутствует, если сумма квадратов выборочных значений r_k превышает порог [ср. (5.86) и (5.87)]

$$\sum_{k=1}^{N} r_k^2 \geqslant \gamma^2, \tag{5.132}$$

где γ — та же константа, что и в (5.87), вычисляемая по формуле (5.88) при очевидной замене $\frac{N_c}{N_0}$ на $\left(\frac{\sigma_c}{\sigma}\right)^2$.

Так как r_k^2 — независимые экспоненциально распределенные случайные величины, то, используя результат, указанный в задаче 3.16 первой книги, находим, что $\sum_{k=1}^{N} r_k^2$ имеет χ_{2N}^2-распределение. Если учесть, что размер выборки значений огибающей процесса вдвое меньше размера выборки самого процесса (см. § 5.1), то отсюда следует, что формулы (5.90) и (5.91) для вероятностей ложных тревог и пропусков сигнала *сохраняются без изменения в случае оптимального амплитудного метода обнаружения*. Таким образом, оптимальные процедуры обнаружения стохастического сигнала до и после детектирования совпадают. Этого совпадения следовало ожидать. Действительно, в рассматриваемом случае функции правдоподобия для выборок фаз, когда сигнал есть и когда его нет, не *отличаются друг от друга*:

$$W_N(\vartheta_1, \ldots, \vartheta_N \mid H_0) = W_N(\vartheta_1, \ldots, \vartheta_N \mid H_1) = \left(\frac{1}{2\pi}\right)^N,$$

так как одномерная плотность вероятности фазы стационарного нормального случайного процесса не зависит от энергетического спектра (дисперсии) процесса и всегда равна $\frac{1}{2\pi}$, $|\vartheta_i| \leqslant \pi$. Отношение правдоподобия в этом случае тождественно равно единице. Отсюда следует также, что в рассматриваемом случае по независимым выборкам фазы невозможно отличить сумму сигнала и шума от чистого шума, т. е. решить задачу обнаружения стохастического сигнала.

Для осуществления фазового метода обнаружения стохастического сигнала необходимо иметь независимые выборки разности фаз *) $\Delta\vartheta_1, \ldots, \Delta\vartheta_N$, где $\Delta\vartheta_i = \vartheta(t_i + \tau) - \vartheta(t_i)$. Используя формулу (8.82) в первой книге, находим функцию правдоподобия выборки разности фаз для шума:

$$W_N(\Delta\vartheta_1, \ldots, \Delta\vartheta_N \mid H_0) = \left(\frac{1-R_0^2}{2\pi}\right)^N \times \prod_{k=1}^{N} \left[\frac{1}{1-y_k^2} + \frac{\frac{\pi}{2} + \arcsin y_k}{(1-y_k^2)^{3/2}} y_k\right], \quad (5.133)$$

$$|\Delta\vartheta_k| \leqslant \pi, \quad k = 1, \ldots, N,$$

*) Это имеет место, например, в задачах разнесенного приема.

где

$$y_k = R_0(\tau)\cos[\vartheta(t_k+\tau) - \vartheta(t_k)]; \qquad (5.134)$$

$$R_0^2(\tau) = R_{c0}^2(\tau) + R_{s0}^2(\tau); \qquad (5.134')$$

$$R_{c0}(\tau) = \frac{1}{\sigma^2}\int_0^\infty F_{ш}(\omega)\cos(\omega-\omega_0)\tau\, d\omega; \qquad (5.135)$$

$$R_{s0}(\tau) = \frac{1}{\sigma^2}\int_0^\infty F_{ш}(\omega)\sin(\omega-\omega_0)\tau\, d\omega; \qquad (5.135')$$

$F_{ш}(\omega)$, σ^2 — энергетический спектр и дисперсия шума.

Функция правдоподобия суммы сигнала и шума равна

$$W_N(\Delta\vartheta_1, \ldots, \Delta\vartheta_N \mid H_1) = \left(\frac{1-R_1^2}{2\pi}\right)^N \times$$

$$\times \prod_{k=1}^{N}\left[\frac{1}{1-z_k^2} + \frac{\frac{\pi}{2} + \arcsin z_k}{(1-z_k^2)^{3/2}} z_k\right], \qquad (5.136)$$

$$|\Delta\vartheta_k| \leqslant \pi, \quad k = 1, \ldots, N,$$

где

168

$$z_k = R_1(\tau)\cos[\vartheta(t_k+\tau) - \vartheta(t_k)]; \qquad (5.137)$$

$$R_1^2(\tau) = R_{c1}^2(\tau) + R_{s1}^2(\tau); \qquad (5.137')$$

$$R_{c1}(\tau) = \frac{1}{\sigma_c^2+\sigma^2}\int_0^\infty [F_c(\omega) + F_{ш}(\omega)]\cos(\omega-\omega_0)\tau\, d\omega; \qquad (5.138)$$

$$R_{s1}(\tau) = \frac{1}{\sigma_c^2+\sigma^2}\int_0^\infty [F_c(\omega) + F_{ш}(\omega)]\sin(\omega-\omega_0)\tau\, d\omega; \qquad (5.138')$$

$F_c(\omega)$, σ_c^2 — энергетический спектр и дисперсия сигнала.

Отношение правдоподобия равно

$$l(\Delta\vartheta_1, \ldots, \Delta\vartheta_N) = \left(\frac{1-R_1^2}{1-R_0^2}\right)^N \times$$

$$\times \prod_{k=1}^{N}\left[\frac{1-y_k^2}{1-z_k^2}\,\frac{1+z_k\left(\frac{\pi}{2}+\arcsin z_k\right)(1-z_k^2)^{-\frac{1}{2}}}{1+y_k\left(\frac{\pi}{2}+\arcsin y_k\right)(1-y_k^2)^{-\frac{1}{2}}}\right], \qquad (5.139)$$

$$|\Delta\vartheta_i| \leqslant \pi, \quad i = 1, \ldots, N.$$

Исследование общего случая представляется весьма трудным. Ограничимся случаем, когда выборки шума

$\vartheta(t_k+\tau)$ и $\vartheta(t_k)$ не коррелированы, т. е. когда $R_0 = 0$ [см. (8.81) в первой книге] и когда $\sigma_c^2 \ll \sigma^2$ (*слабый* сигнал). Тогда, обозначая

$$R^2(\tau)=R_c^2(\tau)+R_s^2(\tau), \qquad (5.140)$$

где

$$R_c(\tau)=\frac{1}{\sigma_c^2}\int_0^\infty F_c(\omega)\cos(\omega-\omega_0)\tau\, d\omega, \qquad (5.141)$$

$$R_s(\tau)=\frac{1}{\sigma_c^2}\int_0^\infty F_c(\omega)\sin(\omega-\omega_0)\tau\, d\omega, \qquad (5.141')$$

и пользуясь малостью величины $\left(\frac{\sigma_c}{\sigma}\right)^2$, получаем

$$R_1(\tau)=\frac{\sigma_c^2}{\sigma^2}R(\tau). \qquad (5.142)$$

Полагая в (5.139) $R_0 \equiv 0$, получаем с учетом (5.142)

$$\ln l(\Delta\vartheta_1, \ldots, \Delta\vartheta_N)\approx\frac{\pi R}{2}\left(\frac{\sigma_c}{\sigma}\right)^2\sum_{k=1}^{N}\cos\Delta\vartheta_k -$$

$$-\frac{N\pi^2R^2}{16}\left(\frac{\sigma_c}{\sigma}\right)^4, \quad |\vartheta_i|<\pi, \quad i=1, \ldots, N. \qquad (5.143)$$

Правило решения может быть теперь сформулировано следующим образом: сигнал присутствует, если

$$\sum_{k=1}^{N}\cos\Delta\vartheta_k > \frac{2}{\pi R}\left(\frac{\sigma}{\sigma_c}\right)^2\ln c+\frac{N\pi R}{8}\left(\frac{\sigma_c}{\sigma}\right)^2. \qquad (5.144)$$

Оптимальная процедура обнаружения состоит в сравнении выхода интегратора (сумматора) косинусов разности фаз с заранее установленным порогом.

Если объем выборки большой ($N \gg 1$), то распределение суммы независимых случайных величин $\cos\Delta\vartheta_k$ близко к нормальному со средними и дисперсией, равными (см. § 8.4.4 в первой книге)

$$m_1\left\{\sum_{k=1}^{N}\cos\Delta\vartheta_k \mid H_0\right\}=0, \qquad (5.145)$$

$$m_1\left\{\sum_{k=1}^{N}\cos\Delta\vartheta_k \mid H_1\right\}=\frac{N\pi R}{4}\left(\frac{\sigma_c}{\sigma}\right)^2, \qquad (5.145')$$

$$M_2\left\{\sum_{k=1}^{N}\cos\Delta\vartheta_k \mid H_0\right\}=M_2\left\{\sum_{k=1}^{N}\cos\Delta\vartheta_k \mid H_1\right\}=\frac{N}{2}. \qquad (5.146)$$

Вероятности ложных тревог и пропусков сигнала при этом равны *)

$$\alpha = 1 - F\left[\frac{\sqrt{2}\ln c}{\pi R\sqrt{N}}\left(\frac{\sigma}{\sigma_c}\right)^2 + \frac{\sqrt{2}\pi R}{8}\left(\frac{\sigma_c}{\sigma}\right)^2\sqrt{N}\right], \tag{5.147}$$

$$\beta = F\left[\frac{\sqrt{2}\ln c}{\pi R\sqrt{N}}\left(\frac{\sigma}{\sigma_c}\right)^2 - \frac{\sqrt{2}\pi R}{8}\left(\frac{\sigma_c}{\sigma}\right)^2\sqrt{N}\right]. \tag{5.148}$$

5.4. РАЗЛИЧЕНИЕ СИГНАЛОВ

5.4.1. Вводные замечания. До сих пор рассматривалась задача обнаружения сигнала по некоторой реализации процесса на входе приемного устройства, который мог быть только шумом или суммой сигнала и шума. Наряду с этой задачей во многих приложениях возникают задачи более сложные. Например, наблюдаемый процесс $x(t)$ может быть суммой шума и сигнала одного из заданной совокупности сигналов $s_0(t), \ldots, s_m(t)$, и необходимо по принятой реализации на интервале $(-T, T)$ процесса $x(t)$ принять решение, какой из указанных сигналов в действительности присутствует в наблюдаемом процессе.

Аналогичная задача *различения сигналов* возникает в технике связи при построении устройств декодирования сообщений. В других случаях приходится выяснять, представляет ли наблюдаемый процесс только шум или его аддитивную смесь с одним или несколькими сигналами из заданной совокупности сигналов. Подобные задачи в радиолокации иногда называют задачами *разрешения сигналов*. Указанные задачи в общей постановке относятся к классу многоальтернативных задач выбора решения (см. § 1.4.9). Ниже будут рассмотрены несколько сравнительно простых задач этого класса, решение которых, однако, имеет самостоятельное значение в ряде практических ситуаций.

---

*) Ср. (5.147) и (5.148) с (6) и (6′) в задаче (5.5). Если учесть, что общий размер выборки в рассматриваемом случае равен $2N$, то можно прийти к заключению, что рабочая характеристика обнаружения стохастического сигнала при указанных ограничениях приближается к рабочей характеристике обнаружения детерминированного сигнала амплитудным методом с параметром характеристики, в $\frac{\pi R}{4}$ раз меньшим $\left(\text{при } a = \frac{\sigma_c}{\sigma}\right)$.

5.4.2. Два детерминированных сигнала *). Пусть известно, что наблюдаемый процесс представляет аддитивную смесь нормального стационарного шума и одного из двух детерминированных сигналов, $s_0(t)$ или $s_1(t)$. Задача состоит в проверке простой гипотезы (H_0), что наблюдаемый процесс нормальный со средним $s_0(t)$, против простой альтернативы (H_1), что этот процесс также нормальный, но со средним, равным $s_1(t)$.

Выберем в соответствии с § 3.4 в качестве наблюдаемых некоррелированные координаты

$$x_k = \sqrt{\lambda_k} \int_{-T}^{T} x(t)\, \varphi_k(t)\, dt, \tag{5.149}$$

где $x(t)$ — реализация принятого процесса на интервале наблюдения $(-T, T)$, а λ_k и $\varphi_k(t)$ — собственные числа и собственные функции интегрального уравнения

$$\varphi(t) = \lambda \int_{-T}^{T} B(y-t)\, \varphi(y)\, dy, \tag{5.150}$$

где $B(\tau)$ — известная корреляционная функция шума.

Ограничиваясь первыми N координатами и учитывая (3.62), запишем логарифм отношения правдоподобия для выборки $x_1, \ldots, x_N$ размером N в виде

$$\ln l(x_1, \ldots, x_N) = \frac{1}{2} \sum_{k=1}^{N} (a_k^2 - b_k^2) + \sum_{k=1}^{N} (b_k - a_k)\, x_k, \tag{5.151}$$

где

$$a_k = \sqrt{\lambda_k} \int_{-T}^{T} s_0(t)\, \varphi_k(t)\, dt; \tag{5.152}$$

$$b_k = \sqrt{\lambda_k} \int_{-T}^{T} s_1(t)\, \varphi_k(t)\, dt. \tag{5.152$'$}$$

Правило выбора решения может быть теперь сформулировано следующим образом: присутствует сигнал $s_1(t)$, если для наблюдаемой выборки

$$\sum_{k=1}^{N} (b_k - a_k)\, x_k \geqslant \ln c + \frac{1}{2} \sum_{k=1}^{N} (b_k^2 - a_k^2), \tag{5.153}$$

*) Этот раздел можно рассматривать как обобщение §. 5.2.1, результаты которого получаются как частный случай при $s_0(t) \equiv 0$ и $s_1(t) \equiv s(t)$.

и решение о присутствии сигнала $s_0(t)$, если выполняется неравенство, обратное (5.153).

Таким образом, алгоритм различения сводится к вычислению взвешенной суммы наблюдаемых координат и сравнению ее с заранее установленным порогом

$$K_N = \ln c + \frac{1}{2} \sum_{k=1}^{N} (b_k^2 - a_k^2). \tag{5.154}$$

Этот алгоритм оптимален по любому из рассмотренных в первой главе критериев качества (непоследовательных). Выбор критерия отражается лишь на величине c.

Повторяя рассуждения, приведенные в § 5.2.1, находим условные вероятности ошибок первого и второго рода (вероятности *перепутывания* сигналов):

$$\alpha = 1 - F\left(\frac{\ln c}{d_N} + \frac{1}{2} d_N\right), \tag{5.155}$$

$$\beta = F\left(\frac{\ln c}{d_N} - \frac{1}{2} d_N\right), \tag{5.156}$$

где

$$d_N^2 = \sum_{k=1}^{N} (b_k - a_k)^2. \tag{5.157}$$

Если результаты наблюдений не дискретизируются, а используется вся реализация $x(t)$ на интервале наблюдения, то правило выбора решения формулируется следующим образом [см. (3.81)]: присутствует сигнал $s_1(t)$, если

$$\int_{-T}^{T} V(t)\, x(t)\, dt \geqslant \ln c + \frac{1}{2} \int_{-T}^{T} V(t)\,[s_0(t) + s_1(t)]\, dt = K_T, \tag{5.158}$$

и принимается решение о том, что присутствует сигнал $s_0(t)$, если выполняется неравенство, обратное (5.158). Функция $V(t)$ представляет решение неоднородного линейного интегрального уравнения

$$\int_{-T}^{T} B(t-u)\, V(u)\, du = s_1(t) - s_0(t). \tag{5.159}$$

Вероятности перепутывания сигналов определяются по формулам (5.155) и (5.156), если в них d_N заменить величиной d_T, определяемой из соотношения [см. (3.87) и (3.88)]

$$d_T^2 = \int_{-T}^{T} V(t)\,[s_1(t) - s_0(t)]\,dt. \qquad (5.160)$$

При различении сигналов на фоне аддитивного *белого* шума решение интегрального уравнения (5.159) имеет вид

$$V(t) = \frac{1}{N_0}\,[s_1(t) - s_0(t)], \qquad (5.161)$$

где N_0 — спектральная плотность шума. В этом случае неравенство (5.158) может быть переписано в виде

$$\int_{-T}^{T} [s_1(t) - s_0(t)]\,x(t)dt \geqslant N_0 \ln c + \frac{1}{2}\int_{-T}^{T} [s_1^2(t) - s_0^2(t)]\,dt = N_0 \ln c + \frac{E_1 - E_0}{2}, \qquad (5.162)$$

где E_1 и E_0 — энергии сигналов $s_1(t)$ и $s_0(t)$ на интервале наблюдения.

Неравенство (5.162) может быть путем элементарных алгебраических преобразований приведено к виду

$$\int_{-T}^{T} [s_0(t) - x(t)]^2\,dt - \int_{-T}^{T} [s_1(t) - x(t)]^2\,dt \geqslant N_0 \ln c. \qquad (5.162')$$

Интегралы в (5.162′) могут быть названы соответственно *расстояниями* наблюдаемой реализации $x(t)$ до сигналов $s_0(t)$ и $s_1(t)$. Правило выбора решения предписывает вычисление разности этих расстояний и сравнение ее с порогом, зависящим от критерия и спектральной плотности белого шума. Для критерия максимального правдоподобия ($c = 1$) принимается решение о наличии того из двух сигналов, который *ближе* к наблюдаемой реализации. Это — правило, согласно которому строится так называемый идеальный (по Котельникову) приемник. Рабочая характеристика такого приемника определяется величиной [см. (5.160)]

$$d_T^2 = \frac{1}{N_0}\int_{-T}^{T} [s_1(t) - s_0(t)]^2\,dt, \qquad (5.163)$$

т. е. отношением расстояния между сигналами к спектральной плотности шума. Если найдена величина d_T^2 из (5.163), то для расчетов могут быть использованы кривые, приведенные на рис. 25 [см. (5.155) и (5.156)].

Заметим, что для *ортогональных* сигналов, для которых

$$\int_{-T}^{T} s_1(t)\, s_0(t)\, dt \equiv 0,$$

величина d_T^2 равна $\frac{E_1+E_2}{N_0}$, т. е. отношению суммарной энергии сигналов к спектральной плотности шума $\left(\text{при } E_1=E_2=E \text{ в этом случае } d_T^2=\frac{2E}{N_0}\right)$. Максимально возможное значение d_T^2, равное $4E/N_0$, реализуется при $s_1(t)=-s_0(t)$.

Из (5.162) следует, что оптимальный приемник для различения двух детерминированных сигналов на фоне аддитивного нормального белого шума состоит из *двух согласованных фильтров* (или коррелометров, см. § 4.3.4) с импульсными переходными функциями

$$h_0^*(\tau)=s_0(T-\tau), \quad |\tau|\leqslant T, \qquad (5.164)$$

$$h_1^*(\tau)=s_1(T-\tau), \quad |\tau|\leqslant T, \qquad (5.164')$$

и следующими за ними устройствами стробирования, вычитания и отсечки. Для сигналов, обладающих одинаковой энергией ($E_1 = E_2$), при использовании критерия максимального правдоподобия последние два устройства заменяются устройством сравнения, так как в этом случае правило (5.162) преобразуется к виду: присутствует сигнал $s_1(t)$, если

$$\int_{-T}^{T} s_1(t)\, x(t)\, dt \geqslant \int_{-T}^{T} s_0(t)\, x(t)\, dt. \qquad (5.165)$$

5.4.3. Два сигнала с неизвестными амплитудами. Предположим теперь, что сигналы известны лишь с точностью до амплитудного множителя A. Задача различения сигналов состоит теперь в проверке *сложной* гипотезы (H_0), что наблюдаемый процесс нормальный со средним $As_0(t)$, против *сложной* альтернативы (H_1), что среднее значение этого нормального процесса равно $As_1(t)$.

Пусть нет никакой априорной информации о параметре A. Используя результаты, приведенные в § 3.5.3, нетрудно убедиться, что логарифм функционала отношения правдоподобия может быть записан в виде

$$\ln l\,[x(t)] = A \int_{-T}^{T} V(t)\,x(t)\,dt -$$

$$-\frac{A^2}{2} \int_{-T}^{T} V(t)\,[s_0(t) + s_1(t)]\,dt, \qquad (5.166)$$

где $V(t)$ — решение интегрального уравнения (5.159).

Следовательно, при $A > 0$ существует следующее равномерно наиболее мощное правило: отвергается гипотеза H_0, т. е. утверждается, что присутствует сигнал $As_1(t)$, если для наблюдаемой реализации $x(t)$

$$\int_{-T}^{T} V(t)\,x(t)\,dt \geqslant K_T. \qquad (5.167)$$

Порог K_T определяется заданной вероятностью α ошибки первого рода:

$$K_T = d_T x_\alpha, \qquad (5.167')$$

где x_α — процентная точка нормального распределения, а d_T находится из (5.160).

Если $A < 0$, то знак неравенства в (5.167) изменяется, а величина K_T заменяется на $-K_T$.

Равномерно наиболее мощное *несмещенное* правило предписывает отвергнуть гипотезу H_0, если

$$\left| \int_{-T}^{T} V(t)\,x(t)\,dt \right| \geqslant K_T. \qquad (5.168)$$

Таким образом, оптимальное (по критерию Неймана — Пирсона) устройство различения двух сигналов с неизвестными амплитудами на фоне аддитивного нормального шума в основном такое же, как и при полностью известных сигналах.

В частности, при $s_0(t) \equiv 0$ из (5.167) следует, что оптимальное правило обнаружения детерминированного

сигнала $As(t)$ с неизвестной амплитудой A сводится к сравнению величины $\int_{-T}^{T} V(t)\,x(t)\,dt$ с порогом, причем $V(t)$ определяется из уравнения (5.22) (в которое переходит (5.159) при $s_1(t) \equiv s(t)$ и $s_0(t) \equiv 0$). А это и предписывается правилом выбора решения (5.21) при обнаружении детерминированного сигнала $s(t)$. Вероятностные характеристики обнаружения будут, конечно, зависеть от величины A, так как параметр d_T [см. (5.23)] в этом случае будет равен

$$d_T^2 = A^2 \int_{-T}^{T} V(t)\,s(t)\,dt.$$

Другой критерий, не связанный с априорной информацией о неизвестном параметре, — критерий максимального правдоподобия (см. § 1.4.3). Как следует из (1.116), в этом случае необходимо найти сначала максимумы функций правдоподобия при двух гипотезах и сравнить их отношение с единицей. Выбирая в качестве наблюдаемых координат величины x_k, определяемые согласно (5.149), и сохраняя обозначения (5.152), (5.152′), запишем функции правдоподобия выборки $x_1, \ldots, x_N$ при двух гипотезах *):

$$W_N(x_1, \ldots, x_N \mid H_0) = \frac{1}{(2\pi)^{N/2}} \exp\left[-\frac{1}{2}\sum_{k=1}^{N}(x_k - Aa_k)^2\right], \tag{5.169}$$

$$W_N(x_1, \ldots, x_N \mid H_1) = \frac{1}{(2\pi)^{N/2}} \exp\left[-\frac{1}{2}(x_k - Ab_k)^2\right]. \tag{5.169′}$$

Учитывая, что максимумы функции $f(A)$ и $\ln f(A)$ достигаются при одном и том же значении A, нетрудно показать, что максимум функции (5.169) достигается при

$$\hat{A}_0 = \frac{\sum_{k=1}^{N} a_k x_k}{\sum_{k=1}^{N} a_k^2}, \tag{5.170}$$

*) Напомним, в § 5.2.1 было показано, что $M_2\{x_i\} = 1$.

а функции (5.169′) — при

$$\hat{A}_1 = \frac{\sum_{k=1}^{N} b_k x_k}{\sum_{k=1}^{N} b_k^2} . \qquad (5.170')$$

Подставляя вместо A в (5.169) величину $\hat{A}_0$, а в (5.169′) — величину $\hat{A}_1$, получим

$$\ln \frac{\max\limits_{A} W_N (x_1, \ldots, x_N \mid H_1)}{\max\limits_{A} W_N (x_1, \ldots, x_N \mid H_0)} =$$

$$= \frac{1}{2d_{N1}^2} \left(\sum_{k=1}^{N} b_k x_k \right)^2 - \frac{1}{2d_{N0}^2} \left(\sum_{k=1}^{N} a_k x_k \right)^2, \qquad (5.171)$$

где

$$d_{N0}^2 = \sum_{k=1}^{N} a_k^2; \qquad (5.172)$$

$$d_{N1}^2 = \sum_{k=1}^{N} b_k^2. \qquad (5.172')$$

Правило выбора решения по критерию максимального правдоподобия может быть сформулировано так: присутствует сигнал $As_1(t)$, если

$$\frac{1}{d_{N1}^2} \left(\sum_{k=1}^{N} b_k x_k \right)^2 \geqslant \frac{1}{d_{N0}^2} \left(\sum_{k=1}^{N} a_k x_k \right)^2 , \qquad (5.173)$$

и сигнал $As_0(t)$, если выполняется неравенство, противоположное (5.173).

В том случае, когда энергии сигналов совпадают ($d_{N0} = d_{N1}$), правило (5.173) переходит в (5.153) при $c = 1$.

Используя результаты § 3.4.4, нетрудно правило (5.173) переписать в аналоговой форме:

$$\frac{1}{d_{T1}^2} \left(\int_{-T}^{T} V_1(t)\, x(t)\, dt \right)^2 \geqslant \frac{1}{d_{T0}^2} \left(\int_{-T}^{T} V_0(t)\, x(t)\, dt \right)^2 , \qquad (5.174)$$

где

$$d_{T1}^2 = \int_{-T}^{T} V_1(t) s_1(t) dt; \qquad (5.175)$$

$$d_{T0}^2 = \int_{-T}^{T} V_0(t) s_0(t) dt, \qquad (5.175')$$

а $V_1(t)$ и $V_0(t)$ — решения интегральных уравнений

$$\int_{-T}^{T} B(t-u) V_1(u) du = s_1(t), \quad |t| \leqslant T, \qquad (5.176)$$

$$\int_{-T}^{T} B(t-u) V_0(t) du = s_0(t), \quad |t| \leqslant T. \qquad (5.176')$$

Оптимальный приемник состоит из двух согласованных фильтров с импульсными переходными функциями

$$h_0^*(\tau) = \frac{1}{d_{T0}} V_0(T-\tau), \quad |\tau| \leqslant T, \qquad (5.177)$$

$$h_1^*(\tau) = \frac{1}{d_{T1}} V_1(T-\tau), \quad |\tau| \leqslant T, \qquad (5.177')$$

и следующими за ними устройствами стробирования, возведения в квадрат и сравнения.

При различении сигналов на фоне аддитивного белого шума решения интегральных уравнений (5.177), (5.177′) имеют вид

$$V_0(t) = \frac{1}{N_0} s_0(t), \quad V_1(t) = \frac{1}{N_0} s_1(t),$$

и из (5.174) следует правило: присутствует сигнал $As_1(t)$, если

$$\frac{1}{E_1} \left(\int_{-T}^{T} s_1(t) x(t) dt \right)^2 \geqslant \frac{1}{E_0} \left(\int_{-T}^{T} s_0(t) x(t) dt \right)^2, \qquad (5.178)$$

где E_0 и E_1 — энергии сигналов $s_0(t)$ и $s_1(t)$ на интервале наблюдения. При $E_0 = E_1$ неравенство (5.178) совпадает, конечно, с (5.165).

Если задано априорное распределение $w_1(A)$ параметра A, то может быть использован байесовский критерий мини-

мального среднего риска. В этом случае в соответствии с § 1.4.2 вычисляется отношение усредненных функций правдоподобия [см. (5.169) и (5.169′)]

$$\Lambda(x_1, \ldots, x_N) = \frac{\int\limits_{-\infty}^{\infty} w_1(A) \exp\left[-\frac{1}{2}\sum\limits_{k=1}^{N}(x_k - Ab_k)^2\right] dA}{\int\limits_{-\infty}^{\infty} w_1(A) \exp\left[-\frac{1}{2}\sum\limits_{k=1}^{N}(x_k - Aa_k)^2\right] dA} \tag{5.179}$$

и в зависимости от того, превышает или нет эта величина порог c (определяемый априорными вероятностями появления сигналов и платами за решения), принимается решение о присутствии сигнала $As_1(t)$ или $As_0(t)$.

В качестве примера рассмотрим случай, когда априорное распределение амплитуды A нормальное со средним A_0 и дисперсией σ_A^2, т. е.

$$w_1(A) = \frac{1}{\sqrt{2\pi\sigma_A^2}} \exp\left[-\frac{(A - A_0)^2}{2\sigma_A^2}\right]. \tag{5.180}$$

Подставляя (5.180) в (5.179) и вычисляя интегралы в числителе и знаменателе, получаем

$$\ln\Lambda(x_1, \ldots, x_N) = \frac{1}{2}\ln\frac{1+\sigma_A^2 d_{N1}^2}{1+\sigma_A^2 d_{N0}^2} + \frac{A_0^2}{2}\times$$
$$\times\left(\frac{d_{N0}^2}{1+\sigma_A^2 d_{N0}^2} - \frac{d_{N1}^2}{1+\sigma_A^2 d_{N1}^2}\right) + A_0\sum_{k=1}^{N}\left(\frac{b_k}{1+\sigma_A^2 d_{N1}^2} -\right.$$
$$\left. - \frac{a_k}{1+\sigma_A^2 d_{N0}^2}\right)x_k + \frac{\sigma_A^2}{2}\left[\frac{1}{1+\sigma_A^2 d_{N1}^2}\left(\sum_{k=1}^{N} b_k x_k\right)^2 -\right.$$
$$\left. - \frac{1}{1+\sigma_A^2 d_{N0}^2}\left(\sum_{k=1}^{N} a_k x_k\right)^2\right], \tag{5.181}$$

где d_{N0}^2 и d_{N1}^2 определяются согласно (5.172) и (5.172′).

Байесовское правило выбора решения формулируется так: присутствует сигнал $As_1(t)$, если

$$A_0 \sum_{k=1}^{N} \left(\frac{b_k}{1+\sigma_A^2 d_{N1}^2} - \frac{a_k}{1+\sigma_A^2 d_{N0}^2} \right) x_k + \\ + \frac{\sigma_A^2}{2} \left[\frac{1}{1+\sigma_A^2 d_{N1}^2} \left(\sum_{k=1}^{N} b_k x_k \right)^2 - \\ - \frac{1}{1+\sigma_A^2 d_{N0}^2} \left(\sum_{k=1}^{N} a_k x_k \right)^2 \right] \geqslant K_N, \qquad (5.182)$$

где

$$K_N = \ln c + \frac{A_0^2}{2} \left(\frac{d_{N1}^2}{1+\sigma_A^2 d_{N1}^2} - \frac{d_{N0}^2}{1+\sigma_A^2 d_{N0}^2} \right) + \\ + \frac{1}{2} \ln \frac{1+\sigma_A^2 d_{N0}^2}{1+\sigma_A^2 d_{N1}^2} . \qquad (5.182')$$

При $\sigma_A \to 0$ функция $w_1(A) \to \delta(A - A_0)$, и неравенство (5.182) преобразуется к виду

$$A_0 \sum_{k=1}^{N} (b_k - a_k) x_k \geqslant \ln c + \frac{A_0^2}{2} (d_{N1}^2 - d_{N0}^2), \qquad (5.183)$$

и, следовательно, правило выбора решения совпадает с байесовским правилом различения полностью известных сигналов (с очевидным условием $A_0 = 1$).

При $\sigma_A \to \infty$ из (5.182) находим

$$\frac{1}{2d_{N1}^2} \left(\sum_{k=1}^{N} b_k x_k \right)^2 - \frac{1}{2d_{N0}^2} \left(\sum_{k=1}^{N} a_k x_k \right)^2 \geqslant \\ \geqslant \ln c + \frac{1}{2} \ln \frac{d_{N0}^2}{d_{N1}^2} . \qquad (5.184)$$

Основные операции, выполняемые над наблюдаемыми координатами, в этом случае такие же, как и при использовании критерия максимального правдоподобия [см. (5.173)]. Отличие состоит лишь в том, что вместо простого сравнения выходов квадрирующих устройств их разности сравниваются с некоторым заранее фиксированным порогом.

Для сигналов равной энергии $d_{N0} = d_{N1}$ формулы (5.182) — (5.184) значительно упрощаются и могут быть переписаны в виде

$$A_0 \sum_{k=1}^{N} (b_k - a_k) x_k + \frac{\sigma_A^2}{2} \sum_{k=1}^{N} (b_k^2 - a_k^2) x_k^2 \geqslant (1 + \sigma_A^2) \ln c,$$

$$A_0 \sum_{k=1}^{N} (b_k - a_k) x_k \geqslant \ln c \quad \text{при} \quad \sigma_A \to 0,$$

$$\sum_{k=1}^{N} (b_k^2 - a_k^2) x_k^2 \geqslant 2 \ln c \quad \text{при} \quad \sigma_A \to \infty.$$

Заметим, что квазидетерминированный процесс $As(t)$ со случайной амплитудой A используется часто в приложениях в качестве модели сигнала с мультипликативной помехой (стационарной), *медленно изменяющейся* за период наблюдения (или на интервале, равном длительности импульсного сигнала).

5.4.4. Два узкополосных сигнала со случайными фазами*). Рассмотрим задачу различения на фоне аддитивного нормального шума двух квазидетерминированных узкополосных сигналов с заданными законами амплитудной и фазовой модуляции и со случайными равномерно распределенными и независимыми начальными фазами

$$s_0(t) = a_0(t) \cos [\omega_0 t - \psi_{s0}(t) + \varphi_0], \tag{5.185}$$

$$s_1(t) = a_1(t) \cos [\omega_1 t - \psi_{s1}(t) + \varphi_1]. \tag{5.185'}$$

Предположим, что

$$0 < \Delta = \frac{\omega_1 - \omega_0}{2} \ll \omega_0,$$

и выберем при представлении узкополосных процессов через огибающую и фазу в качестве центральной частоту

$$\omega^* = \frac{\omega_0 + \omega_1}{2}.$$

Выполняя все рассуждения в той же последовательности, что и в § 5.2.2, нетрудно получить в рассматриваемом случае следующее выражение логарифма отношения усред-

---

*) Результаты § 5.2.2 являются частным случаем при $s_0(t) \equiv 0$, $s_1(t) \equiv s(t)$ приводимых здесь результатов.

ненных функций правдоподобия от N некоррелированных координат комплексной огибающей [см. (5.42)]:

$$\Lambda(z_1, \ldots, z_N) = \exp\left\{-\frac{1}{2}\sum_{k=1}^{N}[|s_{k1}|^2 - |s_{k0}|^2]\right\} \times$$

$$\times \frac{\int\limits_0^{2\pi} \exp[r_{N1}\cos(\varphi_1 - \psi_{N1})]\, d\varphi_1}{\int\limits_0^{2\pi} \exp[r_{N0}\cos(\varphi_0 - \psi_{N0})]\, d\varphi_0}, \qquad (5.186)$$

где

$$r_{N1} = \left|\sum_{k=1}^{N} z_k \overline{s_{k1}}\right|; \quad \psi_{N1} = \operatorname{arctg}\frac{\operatorname{Im}\sum\limits_{k=1}^{N} z_k\overline{s_{k1}}}{\operatorname{Re}\sum\limits_{k=1}^{N} z_k\overline{s_{k1}}}; \qquad (5.187)$$

$$r_{N0} = \left|\sum_{k=1}^{N} z_k \overline{s_{k0}}\right|; \quad \psi_{N0} = \operatorname{arctg}\frac{\operatorname{Im}\sum\limits_{k=1}^{N} z_k\overline{s_{k0}}}{\operatorname{Re}\sum\limits_{k=1}^{N} z_k\overline{s_{k0}}}. \qquad (5.187')$$

Величины s_{k0} и s_{k1} представляют координаты комплексных огибающих сигналов [см. (5.39‴)], определяемые по формулам

$$s_{k0} = \sqrt{\lambda_k}\int\limits_{-T}^{T} z_{s0}(t)\, e^{-it\Delta}\overline{\varphi_k(t)}\, dt, \qquad (5.188)$$

$$s_{k1} = \sqrt{\lambda_k}\int\limits_{-T}^{T} z_{s1}(t)\, e^{it\Delta}\overline{\varphi_k(t)}\, dt, \qquad (5.188')$$

где

$$z_{s0}(t) = a_0(t)\, e^{-i\psi_{s0}(t)}; \quad z_{s1}(t) = a_1(t)\, e^{-i\psi_{s1}(t)}.$$

Выполняя в (5.186) интегрирование и логарифмируя, получаем

$$\ln\Lambda(z_1, \ldots, z_N) - \ln I_0(r_{N1}) - \ln I_0(r_{N0}) -$$

$$-\frac{1}{2}\sum_{k=1}^{N}[|s_{k1}|^2 - |s_{k0}|^2], \qquad (5.189)$$

где $I_0(r)$ — бесселева функция нулевого порядка от мнимого аргумента.

Из (5.189) следует байесовское правило выбора решения: присутствует сигнал $s_1(t)$ (решение γ_1), если для наблюдаемой выборки

$$\ln I_0(r_{N1}) - \ln I_0(r_{N0}) \geqslant \ln c + \\ + \frac{1}{2}\sum_{k=1}^{N} [|s_{k1}|^2 - |s_{k0}|^2] = K_N, \qquad (5.190)$$

и сигнал $s_0(t)$ (решение γ_0), если выполняется неравенство, обратное (5.190).

Если различение проводится не по дискретной выборке, а по реализации $z(t)$ комплексной огибающей, наблюдаемой на интервале $(-T, T)$, то из (5.190) при $N \to \infty$ находим [см. (5.62)]

$$\ln I_0(r_{T1}) - \ln I_0(r_{T0}) \geqslant \ln c + \frac{d_{T1}^2 - d_{T0}^2}{2} = K_T, \qquad (5.191)$$

где

$$r_{T1} = \left| \int_{-T}^{T} U_1(t)\,\overline{z(t)}\,dt \right|, \qquad (5.192)$$

$$d_{T1}^2 = \int_{-T}^{T} U_1(t)\,\overline{z_{s1}(t)}\,e^{-it\Delta}\,dt, \qquad (5.192')$$

$$r_{T0} = \left| \int_{-T}^{T} U_0(t)\,\overline{z(t)}\,dt \right|, \qquad (5.193)$$

$$d_{T0}^2 - \left| \int_{-T}^{T} U_0(z)\,\overline{z_{s0}(t)}\,e^{it\Delta}\,dt \right| \qquad (5.193')$$

а $U_1(t)$ и $U_0(t)$ — решения неоднородных линейных интегральных уравнений

$$\int_{-T}^{T} B_z(t-u)\,U_1(u)\,du = z_{s1}(t)\,e^{it\Delta}, \quad |t| \leqslant T, \qquad (5.194)$$

$$\int_{-T}^{T} B_z(t-u)\,U_0(u)\,du = z_{s0}(t)\,e^{-it\Delta}, \quad |t| \leqslant T, \qquad (5.194')$$

где $z_{s1}(t)$ и $z_{s0}(t)$ — комплексные огибающие сигналов $s_1(t)$ и $s_0(t)$, а $B_z(\tau)$, как и в (5.35'), — корреляционная функция комплексной огибающей шума.

Случайные величины r_{T1}, r_{T0} в неравенстве (5.191) подчиняются обобщенному релеевскому закону, так как пред-

ставляют модуль плоского вектора, компоненты которого независимы и распределены нормально. Средние и дисперсии этих векторов (комплексных случайных величин) равны с учетом (5.194), (5.194′):

$$m_{10}=m_1\left\{\int_{-T}^{T} U_1(t)\,\overline{z(t)}\,dt\,|\,H_0\right\}=$$

$$=\int_{-T}^{T}\int_{-T}^{T} B_z(t-u)\,U_1(t)\,\overline{U_0(u)}\,dt\,du, \qquad (5.195)$$

$$m_{11}=m_1\left\{\int_{-T}^{T} U_1(t)\,\overline{z(t)}\,dt\,|\,H_1\right\}=$$

$$=\int_{-T}^{T}\int_{-T}^{T} B_z(t-u)\,U_1(t)\,\overline{U_1(u)}\,dt\,du, \qquad (5.195')$$

$$m_{00}=m_1\left\{\int_{-T}^{T} U_0(t)\,\overline{z(t)}\,dt\,|\,H_0\right\}=$$

$$=\int_{-T}^{T}\int_{-T}^{T} B_z(t-u)\,U_0(t)\,\overline{U_0(u)}\,dt\,du, \qquad (5.195'')$$

$$m_{01}=m_1\left\{\int_{-T}^{T} U_0(t)\,\overline{z(t)}\,dt\,|\,H_1\right\}=$$

$$=\int_{-T}^{T}\int_{-T}^{T} B_z(t-u)\,U_0(t)\,\overline{U_1(u)}\,dt\,du, \qquad (5.196)$$

$$M_{21}=M_2\left\{\int_{-T}^{T} U_1(t)\,\overline{z(t)}\,dt\,|\,H_0,\,H_1\right\}=$$

$$=\int_{-T}^{T}\int_{-T}^{T} B_z(t-u)\,U_1(t)\,\overline{U_1(t)}\,dt\,du. \qquad (5.196')$$

$$M_{20}=M_2\left\{\int_{-T}^{T} U_0(t)\,\overline{z(t)}\,dt\,|\,H_0,\,H_1\right\}=$$

$$=\int_{-T}^{T}\int_{-T}^{T} B_z(t-u)\,U_0(t)\,\overline{U_0(u)}\,dt\,du. \qquad (5.196'')$$

В качестве простейшей иллюстрации рассмотрим случай различения на фоне аддитивного *белого* шума сигналов при частотной манипуляции, когда $a_1(t) = a_0(t) = A_0$, $\psi_{s1}(t) = \psi_{s0}(t) = 0$, и примем критерий максимального правдоподобия ($c = 1$). Правило (5.191) выбора решения теперь упрощается: присутствует сигнал $s_1(t) = A_0 \cos(\omega_1 t + \varphi_1)$, если

$$\ln I_0(r_{T1}) \geqslant \ln I_0(r_{T0}), \qquad (5.197)$$

и сигнал $s_0(t) = A_0 \cos(\omega_0 t + \varphi_0)$, если выполняется обратное неравенство. Так как логарифм модифицированной бесселевой функции $\ln I_0(r)$ — монотонная функция аргумента r, то правило (5.197) сводится к выяснению, которая из величин r_{T1} или r_{T0} больше. Таким образом, считаем, что присутствует сигнал $s_1(t)$, если

$$r_{T1} \geqslant r_{T0}. \qquad (5.197')$$

Выпишем явное выражение величин в (5.197'), используя то обстоятельство, что в рассматриваемом случае

$$z_{s1}(t) = z_{s0}(t) \equiv A_0$$

и что для белого шума со спектральной плотностью N_0 решения уравнений (5.194) и (5.194') имеют вид

$$U_0(t) = \frac{A_0}{N_0} e^{-it\Delta}, \quad U_1(t) = \frac{A_0}{N_0} e^{it\Delta}.$$

Из (5.197), (5.192) и (5.193) получаем следующее правило различения: присутствует сигнал частоты ω_1, если

$$\left| \int_{-T}^{T} \overline{z(t)} e^{i\frac{\omega_1-\omega_0}{2}} dt \right| \geqslant \left| \int_{-T}^{T} \overline{z(t)} e^{-i\frac{\omega_1-\omega_0}{2}} \right|. \qquad (5.198)$$

Выражая комплексную огибающую через квадратурные составляющие [см. (5.32')], можно (5.198) переписать в виде

$$\int_{-T}^{T} A(t) \cos \frac{\omega_1-\omega_0}{2} t\, dt \int_{-T}^{T} C(t) \sin \frac{\omega_1-\omega_0}{2} t\, dt \geqslant$$

$$\geqslant \int_{-T}^{T} A(t) \sin \frac{\omega_1-\omega_0}{2} t\, dt \int_{-T}^{T} C(t) \cos \frac{\omega_1-\omega_0}{2} t\, dt. \qquad (5.198')$$

Если учесть, что $A(t)$ и $C(t)$ — медленно меняющиеся функции (см. § 6.2.2 в первой книге), и пренебречь интегралами от произведения этих функций на высокочастотное

заполнение вида $\cos \omega t$ (при $\omega > \omega_0$), то правило (5.198′) можно выразить через реализацию принимаемого процесса

$$x(t) = A(t)\cos\frac{\omega_0+\omega_1}{2}t + C(t)\sin\frac{\omega_0+\omega_1}{2}t.$$

Принимается решение, что присутствует сигнал частоты ω_1, если для наблюдаемой на интервале $(-T, T)$ реализации $x(t)$

$$\left(\int_{-T}^{T} x(t)\cos\omega_1 t\,dt\right)^2 + \left(\int_{-T}^{T} x(t)\sin\omega_1 t\,dt\right)^2 \geqslant$$

$$\geqslant \left(\int_{-T}^{T} x(t)\cos\omega_0 t\,dt\right)^2 + \left(\int_{-T}^{T} x(t)\sin\omega_0 t\,dt\right)^2. \quad (5.199)$$

Оптимальный приемник для различения двух гармонических сигналов, соответствующий правилу (5.199), состоит из фильтров, согласованных с этими сигналами (конечной длительности, частот ω_1 и ω_0; см. задачу 4.6), устройств возведения в квадрат, суммирования и сравнения.

Заметим, что в (5.199) сравниваются квадраты значений огибающих наблюдаемого процесса на выходах двух фильтров, согласованных с сигналами. Поэтому оптимальный приемник для различения двух сигналов можно интерпретировать при помощи двух согласованных с сигналами фильтров, линейных детекторов, выделяющих огибающие процессов, и схемы сравнения.

Найдем выражения для вероятностей ошибок. Для рассматриваемого примера из (5.195) — (5.196″) находим

$$m_{10} = m_{01} - d_T^2\frac{\sin 2\Delta T}{2\Delta T}, \quad (5.200)$$

$$m_{11} = m_{00} - M_{21} = M_{20} = d_T^2, \quad (5.201)$$

где

$$d_T^2 = \frac{2A_0^2 T}{N_0}, \quad (5.202)$$

т. е. равно отношению энергии сигнала к спектральной плотности шума.

Если $2T\Delta = (\omega_1 - \omega_0)T = k\pi$, $k = 1, 2, \ldots$, то сигналы *ортогональны* на интервале $(-T, T)$. Из (5.200) следует, что средние значения компонент нормальных слу-

чайных векторов обращаются при этом в нуль. Это означает, что в соответствующих случаях распределения случайных величин r_{T1} и r_{T0} становятся *релеевскими*. При этом возможно получить выражения вероятностей ошибок в замкнутом виде. Действительно, условная вероятность принять решение, что передавался сигнал $s_1(t)$, когда в действительности передавался $s_0(t)$, равна

$$\alpha = P\{\gamma_1 | H_0\} = P\{r_{T1} \geqslant r_{T0} | H_0\} =$$

$$= \frac{1}{d_T^4} \int_0^\infty x \exp\left(-\frac{x^2 + d_T^4}{2d_T^2}\right) I_0(x) \int_x^\infty y e^{-\frac{y^2}{2d_T^2}} dy\, dx =$$

$$= \frac{1}{d_T^2} \int_0^\infty x \exp\left(-\frac{2x^2 + d_T^4}{2d_T^2}\right) I_0(x)\, dx =$$

$$= \frac{1}{2} e^{-\frac{d_T^2}{4}} \int_0^\infty z\sqrt{2} \exp\left[-\frac{(z\sqrt{2})^2 + \left(\frac{d_T}{\sqrt{2}}\right)^2}{2}\right] \times$$

$$\times I_0\left(z\sqrt{2}\,\frac{d_T}{\sqrt{2}}\right) d(z\sqrt{2}),$$

и так как последний интеграл равен единице, то

$$\alpha = \frac{1}{2} e^{-\frac{d_T^2}{4}}. \tag{5.203}$$

Нетрудно убедиться, что условные вероятности ошибок первого и второго рода равны друг другу, так как

$$\beta = P\{\gamma_0 | H_1\} = P\{r_{T0} > r_{T1} | H_1\} =$$

$$= \frac{1}{d_T^4} \int_0^\infty y e^{-\frac{y^2 + d_T^4}{2d_T^2}} I_0(y) \int_y^\infty x e^{-\frac{x^2}{2d_T^2}} dx\, dy = \alpha.$$

Заметим, что алгоритм оптимального различения двух узкополосных сигналов $A\cos(\omega_0 t + \varphi_0)$ и $A\cos(\omega_1 t + \varphi_1)$ с *неизвестной* амплитудой A и случайными равномерно распределенными фазами не отличается от рассмотренного [см. (5.197)]. От величины A будет зависеть только параметр d_T^2 [см. (5.202)], который определяет вероятность ошибки.

5.4.5. Различение многих сигналов. Дальнейшим обобщением *) является задача различения многих сигналов. Пусть известно, что наблюдаемый процесс представляет аддитивную смесь нормального стационарного шума и одного из $m+1$ детерминированных сигналов $s_0(t)$, $s_1(t)$, ..., $s_m(t)$. Теперь совокупность решений содержит не два, а $m+1$ решений $\gamma_0, \gamma_1, \ldots, \gamma_m$, где γ_k — решение, состоящее в том, что наблюдаемый процесс является суммой шума и сигнала $s_k(t)$. Примем байесовский критерий качества при дополнительном предположении, что платы за правильные решения равны нулю, а платы за ошибочные решения одинаковы. Кроме того, допустим, что передача каждого из сигналов равновероятна **). В соответствии с § 1.4.8 оптимальное правило различения сигналов может быть тогда сформулировано следующим образом: присутствует сигнал $s_k(t)$, $k = 1, \ldots, m$, если для наблюдаемой выборки $x_1, \ldots, x_N$

$$\ln l_k(x_1, \ldots, x_N) \geqslant \ln l_j(x_1, \ldots, x_N) \qquad (5.204)$$

при всех $j \neq k$ и

$$\ln l_k(x_1, \ldots, x_N) \geqslant 0 \qquad (5.204')$$

для данного k; присутствует сигнал $s_0(t)$, если для всех k $(k = 1, \ldots, m)$

$$\ln l_k(x_1, \ldots, x_N) < 0. \qquad (5.204'')$$

В приведенных неравенствах отношение правдоподобия

$$l_k(x_1, \ldots, x_N) = \frac{W_N(x_1, \ldots, x_N \mid s_k)}{W_N(x_1, \ldots, x_N \mid s_0)}, \quad k = 1, \ldots, m. \qquad (5.205)$$

Выбирая в качестве наблюдаемых координат величины x_i, определенные согласно (5.149), и ограничиваясь пер-

---

*) Результаты предыдущих разделов этой главы получаются из приводимых ниже при $m = 1$.

**) Если априорная вероятность сигнала $s_k(t)$ равна p_k, то приводимые ниже результаты оказываются верными после замены

$$l_k(x_1, \ldots, x_N) \text{ на } \frac{p_k}{p_0} l_k(x_1, \ldots, x_N), \quad k = 1, \ldots, m.$$

выми N координатами, представим логарифм k-го отношения правдоподобия в виде [см. (5.151)]

$$\ln l_k(x_1, \ldots, x_N) = \frac{1}{2}\sum_{i=1}^{N}(a_{i0}^2 - a_{ik}^2) +$$

$$+ \sum_{i=1}^{N}(a_{ik} - a_{i0})x_i, \quad k = 1, \ldots, m, \qquad (5.206)$$

где

$$a_{ik} = \sqrt{\lambda_i}\int_{-T}^{T} s_k(t)\,\varphi_i(t)\,dt, \quad k = 0, 1, \ldots, m, \qquad (5.207)$$

и λ_i, $\varphi_i(t)$ — собственные числа и собственные функции интегрального уравнения (5.150), ядро которого совпадает с корреляционной функцией шума.

Из (5.204) — (5.204″) и (5.206) следует оптимальное правило различения: присутствует сигнал $s_k(t)$ (выбирается решение γ_k), если для всех $j \neq k$ ($j = 0, 1, \ldots, m$)

$$\sum_{i=1}^{N}(a_{ik} - a_{ij})x_i \geqslant \frac{1}{2}\sum_{i=1}^{N}(a_{ik}^2 - a_{ij}^2). \qquad (5.208)$$

Если результаты наблюдений не дискретизируются, а используется вся реализация $x(t)$ на интервале наблюдения, то при $N \to \infty$ из (5.208) следует правило [см. (5.158)]: присутствует сигнал $s_k(t)$, если для всех $j \neq k$ ($j = 0, 1, \ldots, m$)

$$\int_{-T}^{T} V_{kj}(t)\,x(t)\,dt \geqslant \frac{1}{2}\int_{-T}^{T} V_{kj}(t)\,[s_k(t) + s_j(t)]\,dt, \qquad (5.209)$$

где $V_{kj}(t)$ — решение неоднородного линейного интегрального уравнения

$$\int_{-T}^{T} B(t-u)\,V_{kj}(u)\,du = s_k(t) - s_j(t), \quad |t| \leqslant T. \qquad (5.210)$$

При различении сигналов на фоне аддитивного белого шума решение интегрального уравнения (5.210) имеет вид

$$V_{kj}(t) = \frac{1}{N_0}[s_k(t) - s_j(t)],$$

где N_0 — спектральная плотность шума. В этом случае (5.209) приводится к виду

$$\int_{-T}^{T} s_k(t)\,x(t)\,dt \geqslant \int_{-T}^{T} s_j(t)\,x(t)\,dt + \frac{E_k - E_j}{2}, \quad j \neq k, \qquad (5.211)$$

где E_k и E_j — энергии сигналов $s_k(t)$ и $s_j(t)$ на интервале наблюдения. Неравенство (5.211) равносильно следующему [см. (5.162′)]:

$$\int_{-T}^{T} [s_j(t) - x(t)]^2\,dt \geqslant \int_{-T}^{T} [s_k(t) - x(t)]^2\,dt, \quad j \neq k, \qquad (5.212)$$

которое можно трактовать так: принимается решение о наличии того сигнала $s_k(t)$, для которого *расстояние* до наблюдаемой реализации меньшее по сравнению с расстояниями до любого другого сигнала.

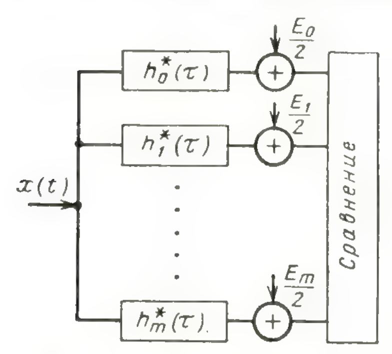


Рис. 28. Блок-схема оптимального устройства различения $m+1$ детерминированных сигналов.

Из (5.211) следует, что оптимальный приемник, предназначенный для различения $m+1$ детерминированных сигналов на фоне аддитивного нормального белого шума, состоит из $m+1$ согласованных фильтров с импульсными переходными функциями

$$h_j^*(\tau) = s_j(T-\tau), \quad |\tau| \leqslant T, \quad j = 0,\ 1,\ \ldots\ m, \qquad (5.213)$$

и следующими за ними устройствами стробирования и сравнения. Для сигналов равной энергии осуществляется перебор значений на выходе согласованных фильтров и принимается решение о наличии того сигнала, для которого это значение максимально (рис. 28).

Рассмотрим также задачу различения на фоне аддитивного нормального шума $m+1$ квазидетерминированных узкополосных сигналов

$$s_k(t) = a_k(t)\cos[\omega_k t - \psi_{sk}(t) + \varphi_k] \qquad (5.214)$$

при условии, что $a_k(t)$ и $\psi_{sk}(t)$ известны точно, а начальные фазы φ_k случайны, распределены равномерно на интервале $(-\pi,\ \pi)$ и независимы.

Предположим, что ω_0 — наименьшая из совокупности частот сигналов, и пусть

$$\Delta_k = \frac{\omega_k - \omega_0}{2} \ll \omega_0.$$

Выберем в качестве центральных частот

$$\omega_k^* = \frac{\omega_k + \omega_0}{2}$$

при представлении через огибающие и фазы узкополосных процессов, содержащих сигналы $s_k(t)$ и $s_0(t)$.

Аналогично (5.189) запишем выражение логарифма k-го отношения усредненных функции правдоподобия:

$$\ln \Lambda_k (z_1, \ldots, z_N) = \ln I_0 (r_{Nk}) - \ln I_0 (r_{N0}) - \\ - \frac{1}{2} \sum_{i=1}^{N} [|s_{ik}|^2 - |s_{i0}|^2], \; k = 1, \ldots, m, \qquad (5.215)$$

где для любого k от 0 до m

$$r_{Nk} = \left| \sum_{i=1}^{N} z_i s_{ik} \right|; \qquad (5.216)$$

$$s_{ik} = \sqrt{\lambda_i} \int_{-T}^{T} z_{sk}(t)\, e^{it\Delta_k} \overline{\varphi_k(t)}\, dt; \qquad (5.217)$$

$$z_{sk}(t) = a_k(t)\, e^{-i\psi_{sk}(t)} \qquad (5.218)$$

и λ_i, $\varphi_i(t)$, как и раньше, — собственные числа и функции однородного линейного интегрального уравнения с ядром, совпадающим с корреляционной функцией $B_z(\tau)$ комплексной огибающей шума [см. (5.35′)].

Из (5.204) — (5.204″) и (5.215) следует правило выбора решения: присутствует сигнал $s_k(t)$ (решение γ_k), если для наблюдаемой выборки $z_1, \ldots, z_N$ и для всех $j \neq k$

$$\ln I_0 (r_{Nk}) - \ln I_0 (r_{Nj}) \geqslant \frac{1}{2} \sum_{i=1}^{N} [|s_{ik}|^2 - |s_{jk}|^2]. \qquad (5.219)$$

Если различение проводится по реализации $z(t)$ комплексной огибающей на интервале $(-T, T)$, то при $N \to \infty$ из (5.219) следует:

$$\ln I_0 (r_{Tk}) - \ln I_0 (r_{Tj}) \geqslant \frac{d_{Tk}^2 - d_{Tj}^2}{2}, \; j \neq k, \qquad (5.220)$$

где для любого j от 0 до m

$$r_{Tj} = \left| \int_{-T}^{T} U_j(t) \overline{z(t)}\, dt \right|, \tag{5.221}$$

$$d_{Tj}^2 = \int_{-T}^{T} U_j(t) \overline{z_{sj}(t)}\, e^{-it\Delta_j}\, dt, \tag{5.222}$$

а $U_j(t)$ — решение неоднородного линейного интегрального уравнения

$$\int_{-T}^{T} B_z(t-y) U_j(y)\, dy = z_{sj}(t)\, e^{it\Delta_j}, \quad |t| < T. \tag{5.223}$$

При различении на фоне аддитивного белого шума со спектральной плотностью N_0 решение этого уравнения имеет вид

$$U_j(t) = \frac{1}{N_0} z_{sj}(t)\, e^{it\Delta_j}, \tag{5.224}$$

а величины r_{Tj} и d_{Tj}^2 равны

$$r_{Tj} = \frac{1}{N_0} \left| \int_{-T}^{T} \overline{z(t)}\, z_{sj}(t)\, e^{it\Delta_j}\, dt \right|, \tag{5.225}$$

$$d_{Tj}^2 = \frac{1}{N_0} \int_{-T}^{T} |z_{sj}(t)|^2\, dt = \frac{E_j}{N_0}, \tag{5.226}$$

где E_j — энергия j-го сигнала.

При различении сигналов равной энергии $E_j = E$ правило выбора решения с учетом монотонности функции $\ln I_0(r)$ может быть сформулировано так: присутствует сигнал $s_k(t)$, если для всех $j \neq k$

$$r_{Tk} \geqslant r_{Tj} \tag{5.227}$$

или

$$\left| \int_{-T}^{T} \overline{z(t)}\, z_{sk}(t)\, e^{it\Delta_k}\, dt \right| \geqslant \left| \int_{-T}^{T} \overline{z(t)}\, z_{sj}(t)\, e^{it\Delta_j}\, dt \right|. \tag{5.227'}$$

Имея в виду, что квадратурные составляющие $z_{sj}(t)$ и $z(t)$ — медленно меняющиеся функции (см. § 6.2.2 в первой книге), и пренебрегая в (5.227′) интегралами от произведений этих функций на высокочастотное заполнение

вида $\cos \omega t$ при $\omega > \omega_0$, можно указать следующее правило различения сигналов (равной энергии) на фоне аддитивного белого шума: принимается решение, что присутствует сигнал $s_k(t)$, если для наблюдаемой реализации $x(t)$ при всех $j \neq k$

$$\left(\int_{-T}^{T} x(t)\, a_k(t) \cos[\omega_k t - \psi_{ks}(t)]\, dt\right)^2 + \left(\int_{-T}^{T} x(t)\, a_k(t) \sin[\omega_k t - \psi_{ks}(t)]\, dt\right)^2 \geqslant$$

$$\geqslant \left(\int_{-T}^{T} x(t)\, a_j(t) \cos[\omega_j t - \psi_{js}(t)]\, dt\right)^2 + \left(\int_{-T}^{T} x(t)\, a_j(t) \sin[\omega_j t - \psi_{js}(t)]\, dt\right)^2. \qquad (5.228)$$

Оптимальный приемник, предназначенный для различения, состоит из набора $2m+2$ фильтров, согласованных с сигналами $s_k(t)$, $|t| \leqslant T$, $k=0, 1, \ldots, m$, и сигналами

$$s_k^*(t) = a_k(t) \sin[\omega_k t - \psi_k(t)], \quad |t| \leqslant T, \; k=0, 1, \ldots, m \qquad (5.229)$$

(которые являются преобразованиями Гильберта от $s_k(t)$; см. приложение VII в первой книге), а также устройство возведения в квадрат, суммирования и выбора наибольшего значения (рис. 29).

Заметим, что в (5.228) сравниваются квадраты значений огибающих наблюдаемого процесса на выходах двух фильтров, согласованных с сигналами $s_k(t)$ и $s_j(t)$. Поэтому оптимальный приемник, предназначенный для различения $m+1$ сигналов, может быть представлен набором только $m+1$ согласованных фильтров, выходные сигналы которых детектируются и поступают далее на схему сравнения, выделяющую наибольшее из $m+1$ значений огибающих (рис. 30).

Нетрудно в рассматриваемом случае определить вероятности перепутывания сигналов, если потребовать, чтобы выполнялись следующие условия *ортогональности*:

при $j \neq k$

$$\int_{-T}^{T} s_k(t)\, s_j(t)\, dt = 0, \qquad (5.230)$$

$$\int_{-T}^{T} s_k^*(t)\, s_j(t)\, dt = 0. \qquad (5.230')$$

Тогда распределение случайной величины r_{Tk} подчиняется релеевскому закону с параметром $d_T^2 = \frac{E}{N_0}$, если

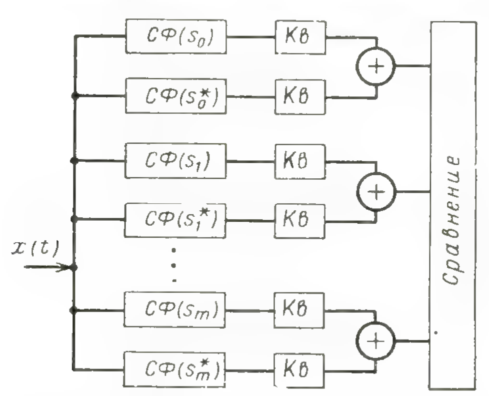


Рис. 29. Блок-схема оптимального устройства различения $m+1$ квазидетерминированных узкополосных сигналов.

гипотеза H_k (присутствует сигнал s_k) не верна, и обобщенному релеевскому распределению с параметрами

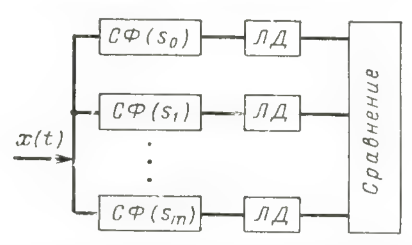


Рис. 30. Блок-схема оптимального устройства различения, эквивалентного изображенному на рис. 29.

(d_T^1, d_T^2), если гипотеза H_k верна. Вероятность α_{kj} перепутать k-й сигнал с j-м, т. е. принять решение, что присутствует $s_j(t)$, когда в действительности послан сигнал

$s_k(t)$, для всех k и j $(k \neq j)$ равна [см. (5.203)]

$$\alpha_{kj} = \frac{1}{2} e^{-\frac{d_T^2}{4}}. \tag{5.231}$$

Вероятность $p_{\text{ош}}$ ошибки, т. е. вероятность назвать *любой* j-й сигнал, когда в действительности присутствует k-й сигнал $(j \neq k)$, находим, используя правило сложения для совместимых событий (см. § 1.1.5 в первой книге):

$$p_{\text{ош}} = 1 - \left(1 - \frac{1}{2} e^{-\frac{d_T^2}{4}}\right)^m. \tag{5.232}$$

Задачи

5.1. Обобщить результаты § 5.2.2 на обнаружение квазидетерминированного сигнала

$$s(t) = Aa(t) \cos [\omega_0 t - \psi_s(t) + \varphi_0)$$

со случайными амплитудой A и фазой φ_0. Исследовать случай, когда фаза не зависит от амплитуды и распределена равномерно. Доказать, что для широкого класса распределений $w_1(A)$ оптимальное байесовское правило обнаружения $s(t)$ на фоне аддитивного нормального шума не зависит от $w_1(A)$ и определяется неравенствами вида (5.44′) или (5.62′). Показать, в частности, что указанное свойство имеет место для релеевского распределения амплитуды. Найти для этого случая (при условии, что $M_2\{A\} = 2$) следующие выражения условной вероятности пропуска сигнала (для критерия Наймана — Пирсона):

$$\beta(A) = e^{-\frac{A^2 d_T^2}{2}} \int_0^K x e^{-\frac{x^2}{2}} I_0(xA\, d_T)\, dx, \tag{1}$$

где порог K определяется заданной вероятностью ложной тревоги α

$$K = \sqrt{2 \ln \frac{1}{\alpha}}, \tag{1′}$$

d_T определяется согласно (5.61). Для дискретной выборки d_T заменяется на d_N [см. (5.49)].

5.2. Для задачи 5.1 при условии, что амплитуда A распределена по релеевскому закону с дисперсией $M_2\{A\} = 2$, найти логарифм усредненного отношения правдоподобия

$$\ln \Lambda(z_1, \ldots, z_N) = -\ln(1 + d_T^2) + \frac{r_T^2}{2(1 + d_T^2)}. \tag{2}$$

Показать, что вероятность правильного обнаружения и вероятность ложной тревоги в этом случае связаны соотношением

$$1 - \beta = \alpha^{\frac{1}{1 + d_T^2}}. \tag{3}$$

В формулах (2) и (3) величины r_T и d_T определяются согласно (5.60) и (5.61) соответственно. При дискретных выборках r_T и d_T заменяются на r_N и d_N [см. (5.42) и (5.49)].

5.3. Показать, что функция распределения суммы квадратов N независимых случайных величин, каждая из которых распределена по обобщенному релеевскому закону

$$w_1(r) = r e^{-\frac{r^2+a^2}{2}} I_0(ar), \quad r > 0,$$

имеет следующий вид:

$$W_1(y) = \frac{1}{2}\left(\frac{y}{a^2}\right)^{\frac{N-1}{2}} e^{-\frac{y+a^2}{2}} I_{N-1}(a\sqrt{y}), \; y > 0. \tag{4}$$

Получить из (4) частный случай при $a=0$ распределения суммы квадратов N независимых релеевских случайных величин (χ^2-распределение с $2N$ степенями свободы)

$$W_1(y) = \frac{1}{2^N (N-1)!} y^{N-1} e^{-y/2}, \qquad y > 0. \tag{4'}$$

5.4. Используя результаты задачи 5.2, найти вероятности ложной тревоги α и пропуска сигнала β при обнаружении по алгоритму (5.101) синусоидального сигнала (отношение амплитуды сигнала к среднеквадратичному значению шума мало). Показать, что при произвольном N

$$\alpha = 1 - \frac{\Gamma\left(N, \frac{2K_N}{a^2}\right)}{\Gamma(N)}, \tag{5}$$

где $\Gamma(N, z)$ — неполная гамма-функция [ср. с (5.90)],

$$\beta = \int_0^{\frac{2\sqrt{K_N}}{a}} x\left(\frac{x}{a}\right)^{N-1} e^{-\frac{x^2+a^2}{2}} I_{N-1}(ax)\, dx =$$

$$= \int_0^{\frac{2\sqrt{K_N}}{a}} x e^{-\frac{x^2+a^2}{2}} I_0(ax)\, dx + e^{-\left(\frac{a^2}{2} + \frac{2K_N}{a^2}\right)} \times$$

$$\times \sum_{k=1}^{N-1} \left(\frac{2\sqrt{K_N}}{a^2}\right)^k I_k(2\sqrt{K_N}), \tag{5'}$$

причем последний интеграл представляет табулированное интегральное обобщенное распределение Релея (см. ссылку [12] к гл. 3 первой книги).

Исследовать критерий максимального правдоподобия, когда

$$K_N = \frac{Na^2}{2}.$$

5.5. Получить следующие асимптотические формулы (при $N \gg 1$) для вероятностей ложных тревог и пропусков сигнала при обнаружении слабого синусоидального сигнала с известной начальной фазой (отношение a амплитуды сигнала к среднеквадратическому значению шума мало):

$$\alpha \sim 1 - F\left(\frac{2\ln c}{a^2\sqrt{N}} + \sqrt{N}\frac{a^2}{2}\right), \tag{6}$$

$$\beta \sim F\left(\frac{2\ln c}{a^2\sqrt{N}} - \sqrt{N}\frac{a^2}{2}\right). \tag{6'}$$

5.6. Показать, что для синусоидального сигнала постоянной амплитуды A_0 при $\frac{A_0}{\sigma} \ll 1$ оптимальное приемное устройство, использующее фазовый метод обнаружения, состоит из идеального ограничителя и интегратора, выход которого сравнивается с заранее установленным порогом:

$$\sum_{k=1}^{N} \cos\vartheta_k > \sqrt{\frac{2}{\pi}}\frac{\sigma}{A_0}\ln c + \frac{N}{4}\sqrt{\frac{\pi}{2}}\frac{A_0}{\sigma} \tag{7}$$

или приближенно при $\Delta T \gg 1$

$$\frac{\Delta}{\pi}\int_{-\pi}^{\pi} \cos\vartheta(t)\,dt \gg \sqrt{\frac{2}{\pi}}\frac{\sigma}{A_0}\ln c + \frac{N}{4}\sqrt{\frac{\pi}{2}}\frac{A_0}{\sigma}. \tag{8}$$

Получить следующие формулы для вероятностей ложных тревог и пропусков сигнала:

$$\alpha = 1 - F\left(\frac{\ln c}{\sqrt{\pi N}}\frac{2\sigma}{A_0} + \sqrt{\pi N}\frac{A_0}{4\sigma}\right), \tag{9}$$

$$\beta - F\left(\frac{\ln c}{\sqrt{\pi N}}\frac{2\sigma}{A_0} - \sqrt{\pi N}\frac{A_0}{4\sigma}\right). \tag{10}$$

5.7. Показать, что для синусоидального сигнала постоянной амплитуды A_0 с известной начальной фазой при $\frac{A_0}{\sigma} \gg 1$ оптимальное приемное устройство, использующее фазовый метод обнаружения, состоит из фазового детектора, выделяющего квадрат фазы и интегратора. Доказать, что правило выбора решения (критерий максимального правдоподобия) определяется неравенствами: сигнал присутствует, если

$$\sum_{k=1}^{N} \vartheta_k^2 \leqslant \frac{2N\sigma^2}{A_0^2}\ln\left(\sqrt{2\pi}\frac{A_0}{\sigma}\right) \tag{11}$$

или приближенно при $\Delta T \gg 1$

$$\int_{-T}^{T} \vartheta^2(t)\,dt \leqslant \frac{2\sigma^2}{\Delta A_0^2} 2\Delta T \ln\left(2\pi \frac{A_0}{\sigma}\right). \qquad (12)$$

Получить следующие асимптотические формулы для вероятностей ложных тревог и пропусков сигнала

$$\alpha \sim F\left\{\frac{\sqrt{45}}{4}\sqrt{N}\left[1 - \frac{2\ln\left(\sqrt{2\pi}\,\frac{A_0}{\sigma}\right)}{\pi^2 A_0^2/\sigma^2}\right]\right\}, \qquad (13)$$

$$\beta \sim 1 - F\left[\sqrt{2N}\ln\left(\sqrt{2\pi}\,\frac{A_0}{\sigma}\right)\right]. \qquad (14)$$

5.8. Показать, что вероятность обнаружения синусоидального сигнала постоянной амплитуды A_0 при $\frac{A_0}{\sigma} \gg 1$, когда используется алгоритм (5.128′) и критерий Неймана–Пирсона, может быть определена по формуле

$$1 - \beta = \frac{1}{\Gamma\left(\frac{N-1}{2}\right)}\Gamma\left[\frac{N-1}{2}, \sqrt{\frac{2}{N-1}}\left(\frac{A_0}{\sigma}\right)^2 (N-K)\right], \qquad (15)$$

где $\Gamma(m, \lambda)$ — неполная гамма-функция и порог K определяется вероятностью ложных тревог.

5.9. Проверяется гипотеза H_0 о том, что реализация $x(t)$, $|t| \leqslant T$, принадлежит стохастическому сигналу с корреляционной функцией $B_0(u, v)$, против альтернативы H_1, что эта реализация принадлежит стохастическому сигналу с корреляционной функцией $B_1(u, v)$. Распределения сигналов нормальные. Показать, что оптимальное устройство различения двух указанных стохастических сигналов не отличается от изображенного на рис. 27, а импульсная переходная функция $h(u, v)$ линейного фильтра с переменными параметрами определяется из интегрального уравнения (см. § 3.5.7, а также [21])

$$\int_{-T}^{T}\int_{-T}^{T} B_0(t, u) B_1(v, y) h(u, v)\,du\,dv = B_1(t, y) - B_0(t, y), \qquad (16)$$

$$|t| \leqslant T, \quad |y| \leqslant T.$$

5.10. Показать, что вероятность ошибки при различении $m+1$ детерминированных ортогональных сигналов $s_k(t)$, $k = 0, \ldots, m$, равной энергии E на фоне аддитивного нормального белого шума с интенсивностью N_0 равна

$$p_{\text{ош}} = 1 - \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} F^m\left(x + \sqrt{\frac{2E}{N_0}}\right)dx. \qquad (17)$$

5.11. Заменить в задаче 5.10 условие ортогональности условием

$$\int_{-T}^{T} s_k(t)\, s_j(t)\, dt = \lambda E, \ \lambda \leqslant 1, \tag{18}$$

и показать, что при этом

$$p_{\text{ош}} = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} F^m \left(x + \sqrt{\frac{2E(1-\lambda)}{N_0}} \right) dx. \tag{19}$$

Доказать, что в (18) временной коэффициент взаимной корреляции сигналов λ не может быть меньше чем $-\frac{1}{m}$ (и, следовательно, минимально возможное значение λ, соответствующее наименьшей вероятности ошибки, равно $-\frac{1}{m}$).

ЛИТЕРАТУРА

МОНОГРАФИИ

1. Вайнштейн Л. А., Зубаков В. Д. Выделение сигналов на фоне случайных помех. Изд-во «Советское радио», 1960, гл. I—III.
2. Давенпорт Г. В., Рут Б. Л. Введение в теорию случайных сигналов и шумов. Пер. с англ., под ред. Р. Л. Добрушина. Изд-во иностранной литературы, 1960, гл. 14.
3. Гуткин Л. С. Теория оптимальных методов радиоприема при флуктуационных помехах. Госэнергоиздат, 1961, гл. 9, 10.
4. Котельников В. А. Теория потенциальной помехоустойчивости. Госэнергоиздат, 1956.
5. Левин Б. Р. Теория случайных процессов и ее применение в радиотехнике, изд. 2-е. Изд-во «Советское радио», 1960, гл. 11.
6. Миддлтон Д. Введение в статистическую теорию связи. Пер. с англ., под ред. Б. Р. Левина, т. II, Изд-во «Советское радио», 1962, гл. 19, 20.
7. Миддлтон Д. Очерки теории связи. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1966, гл. 2.
8. Финк Л. М. Теория передачи дискретных сообщений. Изд-во «Советское радио», 1963, гл. 4.
9. Харкевич А. А. Борьба с помехами. Физматгиз, 1964, § 8—14.
10. Тихонов В. И. Статистическая радиотехника. Изд-во «Советское радио», 1966, гл. 11.
11. Хелстром К. Статистическая теория обнаружения сигналов. Пер. с англ., под ред. Ю. Б. Кобзарева. Изд-во иностранной литературы, 1963, гл. 5.
12. Ширман Я. Д., Голиков В. Н. Основы теории обнаружения радиолокационных сигналов и измерения их параметров. Изд-во «Советское радио», гл. 4—6.
13. Schwartz M., Bennett W. R., Stein S. Communication system and techniques. Mc Graw Hill, 1966, pt. 1.

14. Selin I. Detection theory. New Jersey, Princeton Univ. Press, 1965.
15. Wald A. Statistical Decision functions, N.Y., 1950.
16. Viterbi J. V., Principles of coherent communication. Mc Graw Hill, N.Y., 1966.

СТАТЬИ

17. Левин Б. Р. Оптимальные фазовые методы обнаружения сигналов. «Радиотехника и электроника», 1960, № 4.
18. Питерсон У., Бердсал Т., Фокс У. Теория обнаружения сигналов. В сборнике «Теория информации и ее приложения». Пер. с англ., под ред. А. А. Харкевича. Физматгиз, 1959.
19. Esposito R., Middleton D., Mullen J. A. Advantages of signals subject to slow Reyleigh Fading. IEEE Trans., 1965, IT-11, № 4.
20. Kelly E. J., Reed I. S. Root W. L. The detection of radar echoes in noise I. J. Sos. Ind. Appl., Math., 1960, v. 8, № 2.
21. Kadota T. T. Optimum reception of binary gaussian signals. BSTJ, 1964, Nov., v. 43.
22. Kadota T. T. Simultaneously orthogonal expansion of two stationary gaussian processes. BSTJ, 1966, Sept., v. 45.
23. Черняк Ю. Б. О линейных свойствах системы широкополосный ограничитель — фильтр. «Радиотехника и электроника», 1962, № 7.

ГЛАВА ШЕСТАЯ

ВЫДЕЛЕНИЕ СИГНАЛОВ НА ФОНЕ ПОМЕХ

6.1. ОЦЕНКИ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ НЕИЗВЕСТНЫХ ПАРАМЕТРОВ СИГНАЛА

6.1.1. Постановка задачи. Переходим к изложению некоторых вопросов, относящихся к проблеме *выделения* сигналов на фоне помех. Как уже отмечалось в начале пятой главы, эта проблема возникает в ситуации, когда априори *известно*, что сигнал содержится в наблюдаемом случайном процессе, представляющем комбинацию сигнала и помехи, но *неизвестны* некоторые характеристики сигнала, которые в закодированном виде содержат полезную информацию. Эти характеристики могут быть неизвестными числами или функциями времени, случайными величинами или случайными процессами. Задача выделения сигнала состоит в определении оптимальных (в смысле принятых критериев качества) процедур построения *оценок* указанных характеристик по результатам наблюдения, которые могут быть представлены в виде выборки конечного размера или непрерывной реализации процесса на входе приемника *).

Как и в предыдущей главе, изложение ограничивается в основном случаем, когда помехи представляют аддитивный нормальный стационарный шум, статистически независимый от сигнала. Вначале предполагается, что наблюдаемый процесс представляет сумму шума $\xi(t)$ с нулевым средним и детерминированного сигнала $s(t; A, \varphi, \vartheta_1, \ldots, \vartheta_M)$, зависящего от неизвестных параметров A, φ, $\vartheta_1, \ldots, \vartheta_M$. Затем рассматриваются байесовские оценки случайных параметров квазидетерминированного сигнала при известном их совместном априорном распределении.

---

*) Заметим, что процесс оценки характеристик сигнала может быть совмещен с обнаружением сигнала (подробнее см. в [11]).

Содержание § 6.1 и 6.2 является детализацией и радиотехнической интерпретацией общих математических результатов, указанных в § 3.6.

6.1.2. Совместные оценки амплитуды и фазы гармонического сигнала. Оценка неизвестной амплитуды a детерминированного сигнала $as(t)$ в аддитивном нормальном шуме была подробно рассмотрена в § 3.6.2.

Так, было показано, что несмещенная эффективная оценка неизвестной амплитуды получается интегрированием реализации процесса $x(t)$ с весом, определяемым корреляционной функцией шума. В случае белого шума эта оценка получается на выходе фильтра, согласованного с сигналом $s(t)$, а отношение дисперсии оценки (минимально возможной) к квадрату оцениваемой амплитуды равно при этом отношению спектральной плотности шума к энергии сигнала [см. (3.177)].

Теперь исследуем простейший случай совместной оценки двух неизвестных параметров — амплитуды A и фазы φ детерминированного гармонического сигнала $s(t) = A\cos(\omega_0 t - \varphi)$ известной частоты ω_0 в аддитивном узкополосном нормальном шуме с известной корреляционной функцией $B(\tau)$. Воспользовавшись комплексными представлениями наблюдаемой реализации и сигнала

$$x(t) = \operatorname{Re} z(t)\, e^{i\omega_0 t}, \tag{6.1}$$

$$s(t;\ A,\ \varphi) = \operatorname{Re} A\, e^{-i\varphi}\, e^{i\omega_0 t}, \tag{6.1'}$$

запишем систему уравнений *максимального правдоподобия*, из которой получается соответствующие этому критерию оценки. В соответствии с (3.161′) при $V(t;\ A,\ \varphi) = Ae^{-i\varphi}\, U(t)$ имеем

$$e^{-i\varphi}\int_{-T}^{T} U(t)\overline{z(t)}\, e^{-i\omega_0 t}\, dt - A\int_{-T}^{T} U(t)\, e^{-i\omega_0 t} dt + {}$$

$$+ e^{i\varphi}\int_{-T}^{T} \overline{U(t)}\, z(t)\, e^{i\omega_0 t}\, dt - A\int_{-T}^{T} \overline{U(t)}\, e^{i\omega_0 t}\, dt = 0, \tag{6.2}$$

$$e^{-i\varphi}\int_{-T}^{T} U(t)\overline{z(t)}\, e^{-i\omega_0 t}\, dt - e^{i\varphi}\int_{-T}^{T} \overline{U(t)}\, \times$$

$$\times z(t)\, e^{i\omega_0 t}\, dt = 0, \tag{6.2'}$$

где $U(t)$ определяется из интегрального уравнения

$$\int_{-T}^{T} B_z(t-y) U(y) dy = e^{i\omega_0 t}, \qquad |t| \leqslant T. \tag{6.3}$$

Складывая (6.2) и (6.2′), получаем уравнение, связывающее оценки $\hat{A}$ и $\hat{\varphi}$:

$$\hat{A} e^{i\hat{\varphi}} = \frac{\int_{-T}^{T} U(t) \overline{z(t)} e^{-i\omega_0 t} dt}{\mathrm{Re} \int_{-T}^{T} U(t) e^{-i\omega_0 t} dt}. \tag{6.4}$$

Из (6.4) непосредственно следуют формулы для оценки неизвестных амплитуды и фазы:

$$\hat{A} = \frac{\left| \int_{-T}^{T} U(t) \overline{z(t)} e^{-i\omega_0 t} dt \right|}{\mathrm{Re} \int_{-T}^{T} U(t) e^{-i\omega_0 t} dt}, \tag{6.5}$$

$$\hat{\varphi} = \mathrm{arctg} \frac{\mathrm{Im}\, Y}{\mathrm{Re}\, Y}, \tag{6.5′}$$

где

$$Y = \int_{-T}^{T} \overline{U(t)} z(t) e^{-i\omega_0 t} dt. \tag{6.5″}$$

Если аддитивный шум *белый* со спектральной плотностью N_0, то $U(t) = \frac{1}{N_0} e^{i\omega_0 t}$ и из (6.4) получаем

$$\hat{A} e^{i\hat{\varphi}} = \frac{1}{2T} \int_{-T}^{T} z(t) dt. \tag{6.6}$$

Из (6.6) находим следующую оценку фазы:

$$\hat{\varphi} = \mathrm{arctg} \frac{\mathrm{Im} \int_{-T}^{T} z(t) dt}{\mathrm{Re} \int_{-T}^{T} z(t) dt} = \mathrm{arctg} \frac{\int_{-T}^{T} c(t) dt}{\int_{-T}^{T} A(t) dt}, \tag{6.7}$$

где $A(t)$ и $C(t)$ — квадратурные составляющие наблюдаемой реализации (см. § 5.2.2). Учитывая, что эти состав-

ляющие представляют медленно меняющиеся по сравнению с $\cos \omega_0 t$ процессы, можно пренебречь интегралами от произведений $A(t)$ и $C(t)$ на высокочастотные функции $\cos \omega t$ при $\omega \geqslant \omega_0$. Тогда оценка (6.7) в зависимости от наблюдаемой реализации $x(t)$ представляется следующим образом:

$$\hat{\varphi} = \operatorname{arctg} \frac{\int\limits_{-T}^{T} x(t) \sin \omega_0 t\, dt}{\int\limits_{-T}^{T} x(t) \cos \omega_0 t\, dt}. \tag{6.8}$$

Оценка амплитуды в этом случае равна

$$\hat{A} = \left| \frac{1}{2T} \int\limits_{-T}^{T} z(t)\, dt \right| =$$

$$= \frac{1}{2T} \sqrt{\left(\int\limits_{-T}^{T} A(t)\, dt \right)^2 + \left(\int\limits_{-T}^{T} C(t)\, dt \right)^2} \tag{6.9}$$

или с учетом замечания, сделанного при выводе (6.8),

$$\hat{A} = \left[\left(\frac{1}{2T} \int\limits_{-T}^{T} x(t) \cos \omega_0 t\, dt \right)^2 + \right.$$

$$\left. + \left(\frac{1}{2T} \int\limits_{-T}^{T} x(t) \sin \omega_0 t\, dt \right)^2 \right]^{\frac{1}{2}}. \tag{6.10}$$

Сравнивая (6.9), (6.10) с (5.67), (5.67′), приходим к выводу, что устройство обнаружения квазидетерминированного гармонического сигнала в белом шуме может быть непосредственно использовано для получения оценки амплитуды сигнала, а его отдельные элементы (интеграторы квадратур) — для оценки фазы [см. (6.7) и (6.8)].

Рассмотрим основные свойства оценок амплитуды и фазы гармонического сигнала, принимаемого на фоне аддитивного белого шума. Прежде всего отметим, что интегралы $\frac{1}{2T} \int\limits_{-T}^{T} A(t)\, dt$ и $\frac{1}{2T} \int\limits_{-T}^{T} C(t)\, dt$ представляют независимые случайные величины, подчиняющиеся нормальному закону распределения с параметрами $\left(A \cos \varphi, \frac{N_0}{2T} \right)$ и $\left(A \sin \varphi, \frac{N_0}{2T} \right)$

соответственно. Таким образом, из (6.7) и (6.9) следует, что оценки $\hat{A}$ и $\hat{\varphi}$ представляют *модуль* и *фазу* плоского вектора с независимыми нормально распределенными компонентами, имеющими одинаковые дисперсии. Отсюда следует, что оценка $\hat{A}$ амплитуды A [см. (6.9)] имеет обобщенное релеевское распределение с параметрами $\left(A, \frac{N_0}{2T}\right)$, причем при $d_T \to \infty$ это распределение *асимптотически нормальное*. Среднее значение оценки $\hat{A}$ равно (см. (3.74) в первой книге)

$$m_1\{\hat{A}\} = \frac{A}{d_T}\sqrt{\frac{\pi}{2}}\left[\left(1+\frac{d_T^2}{2}\right) I_0\left(\frac{d_T^2}{4}\right) + \right.$$

$$\left. + \frac{d_T^2}{2} I_1\left(\frac{d_T^2}{4}\right)\right] e^{-\frac{d_T^2}{4}}, \qquad (6.11)$$

где d_T^2 — отношение энергии сигнала к спектральной плотности шума, т. е.

$$d_T^2 = \frac{2TA^2}{N_0}. \qquad (6.11')$$

Таким образом, оценка максимального правдоподобия амплитуды оказывается *смещенной*.

Если $d_T^2 \gg 1$, то (см. (3.77) в первой книге)

$$m_1\{\hat{A}\} \sim A\left(1+\frac{1}{2d_T^2}\right), \qquad (6.12)$$

и, следовательно, $\hat{A}$ при $d_T \to \infty$ — *асимптотически несмещенная* оценка амплитуды сигнала.

Дисперсия оценки $\hat{A}$ равна (см. (3.75) в первой книге)

$$M_2\{\hat{A}\} = A^2\left(1+\frac{2}{d_T^2}\right) - m_1^2\{\hat{A}\}. \qquad (6.13)$$

Если $d_T^2 \gg 1$, то (см. (3.78) в первой книге)

$$M_2\{\hat{A}\} \sim \frac{A^2}{d_T^2}\left(1-\frac{1}{4d_T^2}\right). \qquad (6.14)$$

Нетрудно определить функцию распределения оценки максимального правдоподобия фазы $\hat{\varphi}$ (так как эта оценка,

как указывалось выше, представляет фазу вектора с независимыми нормальными компонентами)

$$\hat{\varphi} = \operatorname{arctg} \frac{\int\limits_{-T}^{T} C(t)\, dt}{\int\limits_{-T}^{T} A(t)\, dt}. \tag{6.15}$$

Тогда в соответствии с (3.81) из первой книги плотность вероятности $\hat{\varphi}$ равна

$$w_1(\vartheta) = \frac{1}{2\pi} e^{-\frac{d_T^2}{2}} \left\{1 + d_T \sqrt{2\pi} \cos(\vartheta - \varphi)\, e^{\frac{d_T^2 \cos^2(\vartheta-\varphi)}{2}} \times \right.$$
$$\left. \times F\left[d_T \cos(\vartheta - \varphi)\right]\right\}, \quad |\vartheta - \varphi| \leqslant \pi. \tag{6.16}$$

Так как функция $w_1(\vartheta)$ четная относительно $\vartheta - \varphi$, то указанное распределение свидетельствует о том, что оценка максимального правдоподобия представляет *несмещенную* оценку неизвестной фазы гармонического сигнала в аддитивном нормальном шуме.

Дисперсия оценки $\hat{\varphi}$ равна (см. (3.93) в первой книге)

$$M_2\{\hat{\varphi}\} = \frac{\pi^2}{3} + 4\pi \sum_{n=1}^{\infty} (-1)^n \frac{a_n}{n^2}, \tag{6.17}$$

где

$$a_n = \frac{\Gamma\left(1 + \frac{n}{2}\right) d_T^n}{\pi n!\, 2^{n/2}}\, {}_1F_1\left(\frac{n}{2},\ n+1,\ -\frac{d_T^2}{2}\right). \tag{6.17'}$$

Если $d_T^2 \gg 1$, то из (6.16) следует, что распределение $\hat{\varphi}$ асимптотически нормальное с нулевым средним и дисперсией (см. (3.88) в первой книге)

$$M_2\{\hat{\varphi}\} \sim \frac{1}{d_T^2}. \tag{6.18}$$

Определив информационную матрицу относительно оцениваемых параметров A и φ, можно показать, что оценки $\hat{A}$ и $\hat{\varphi}$ максимального правдоподобия асимптотически совместно эффективны (см. задачу 6.3).

Таким образом, в соответствии с общей теорией (см. § 2.3.1) рассмотренные оценки максимального правдоподобия амплитуды и фазы сигнала в аддитивном нормальном белом шуме *асимптотически несмещенны, совместно эффек-*

тивны и *нормальны*, причем оценка фазы несмещенная (при любом конечном d_T), а оценка амплитуды смещенная.

6.1.3. Оценки параметров узкополосного сигнала на фоне аддитивного белого шума. Рассмотрим узкополосный сигнал

$$s(t;\, A,\, \varphi,\, \vartheta_1,\, \ldots,\, \vartheta_M) = A\,\mathrm{Re}\, a(t;\, \vartheta_1,\, \ldots,\, \vartheta_M)\, e^{-i\varphi}\, e^{i\omega_0 t}, \tag{6.19}$$

который кроме амплитуды $A > 0$ и фазы φ содержит еще M неизвестных параметров комплексной огибающей. Требуется определить оценки максимального правдоподобия этих параметров по реализации процесса $x(t) = \mathrm{Re}\, z(t)\, e^{i\omega_0 t}$, представляющего аддитивную смесь указанного сигнала и *нормального белого шума* со спектральной интенсивностью N_0.

Используя (5.54) (с очевидным изменением некоторых обозначений), запишем логарифм функционала отношения правдоподобия, учитывая, что для рассматриваемого случая

$$V(t) = \frac{A}{N_0}\, a(t;\, \vartheta_1,\, \ldots,\, \vartheta_M)\, e^{-i\varphi},$$

в виде

$$\ln l\,[z(t) \mid A,\, \varphi,\, \vartheta_1,\, \ldots,\, \vartheta_M] = \frac{2A}{N_0}\,\mathrm{Re}\, e^{-i\varphi} \int_{-T}^{T} a(t;\, \vartheta_1,\, \ldots,\, \vartheta_M)\, \overline{z(t)}\, dt - \frac{A^2}{N_0} \int_{-T}^{T} |a(t;\, \vartheta_1,\, \ldots,\, \vartheta_M)|^2\, dt. \tag{6.20}$$

Пусть сигналы нормированы таким образом, что при любых $\vartheta_1,\, \ldots,\, \vartheta_M$

$$\frac{1}{2T} \int_{-T}^{T} |a(t;\, \vartheta_1,\, \ldots,\, \vartheta_M)|^2\, dt = 1. \tag{6.21}$$

Условие (6.21) означает постоянство энергии сигнала на интервале наблюдения, так как

$$E = \int_{-T}^{T} s^2(t;\, A,\, \varphi,\, \vartheta_1,\, \ldots,\, \vartheta_M)\, dt = A^2 \int_{-T}^{T} |a(t;\, \vartheta_1,\, \ldots,\, \vartheta_M)|^2 \cos^2 [\omega_0 t + \arg a(t;\, \vartheta_1,\, \ldots,\, \vartheta_M) - \varphi]\, dt \approx A^2 T,$$

где, как обычно, пренебрегается интегралом от произведения медленно меняющейся функции и высокочастотного заполнения.

Учитывая (6.21), перепишем (6.20) в виде

$$\ln l\,[z(t)\,|\,A,\ \varphi,\ \vartheta_1,\ \ldots,\ \vartheta_M] =$$
$$= \frac{1}{2TN_0}\left|\int\limits_{-T}^{T}\overline{a(t;\ \vartheta_1,\ \ldots,\ \vartheta_M)}\,z(t)\,dt\right|^2 -$$
$$-\left|A\sqrt{\frac{2T}{N_0}}\,e^{i\varphi} - \frac{1}{\sqrt{2TN_0}}\int\limits_{-T}^{T}\overline{a(t;\ \vartheta_1,\ \ldots,\ \vartheta_M)}\,z(t)\,dt\right|^2. \tag{6.22}$$

Пусть $\hat{A}$, $\hat{\varphi}$, $\hat{\vartheta}_1$, ..., $\hat{\vartheta}_M$ — оценки максимального правдоподобия. Так как первый член (6.22) положительный и не зависит от A и φ, то максимальное значение логарифма функционала отношения правдоподобия по параметрам A и φ достигается при условии обращения в нуль второго (отрицательного) члена. Отсюда находим соотношение между оценками

$$\hat{A}e^{i\hat{\varphi}} = \frac{1}{2T}\int\limits_{-T}^{T}\overline{\hat{a}(t;\ \vartheta_1,\ \ldots,\ \hat{\vartheta}_M)}\,z(t)\,dt, \tag{6.23}$$

откуда

$$\hat{A} = \left|\frac{1}{2T}\int\limits_{-T}^{T}\overline{a(t;\ \hat{\vartheta}_1,\ \ldots,\ \hat{\vartheta}_M)}\,z(t)\,dt,\right. \tag{6.23'}$$

$$\hat{\varphi} = \operatorname{arctg}\frac{\operatorname{Im}\int\limits_{-T}^{T}\overline{a(t;\ \hat{\vartheta}_1,\ \ldots,\ \hat{\vartheta}_M)}\,z(t)\,dt}{\operatorname{Re}\int\limits_{-T}^{T}\overline{a(t;\ \hat{\vartheta}_1\ \ldots,\ \hat{\vartheta}_M)}\,z(t)\,dt}\,. \tag{6.23''}$$

Заметим, что при $a(t) \equiv 1$ формула (6.23), как и следует ожидать, переходит в (6.6).

Теперь задача состоит в определении максимума максиморума первого слагаемого (6.22), т. е. следующей функции параметров ϑ_1, ... $.\vartheta_M$:

$$Q(\vartheta_1,\ \ldots,\ \vartheta_M) = \frac{1}{2TN_0}\left|\int\limits_{-T}^{T}\overline{a(t;\ \vartheta_1,\ \ldots,\ \vartheta_M)}\,z(t)\,dt\right|^2. \tag{6.24}$$

В рассматриваемой задаче *выделения сигнала* известно, что наблюдаемая реализация представляет сумму сигнала и шума. Комплексная огибающая этой реализации равна

$$z(t) = A_0 a(t;\ \vartheta_1^0,\ \ldots,\ \vartheta_M^0)\, e^{-i\varphi_0} + z_1(t), \tag{6.25}$$

где $z_1(t)$ — комплексный нормальный стационарный случайный процесс с равномерным спектром, а величины A_0, φ_0, $\vartheta_1^0, \ldots, \vartheta_M^0$ представляют истинные значения параметров сигнала, которые необходимо оценить по наблюдаемой реализации $x(t) = \operatorname{Re} z(t)\, e^{i\omega_0 t}$.

Подставляя (6.25) в (6.24), представим функцию $Q(\vartheta_1, \ldots, \vartheta_M)$ в виде

$$Q(\vartheta_1,\ \ldots,\ \vartheta_M) = \frac{2TA_0^2}{N_0}\,[h(\vartheta_1,\ \ldots,\ \vartheta_M) + \\ + 2\operatorname{Re}\overline{\Psi(\vartheta_1,\ \ldots,\ \vartheta_M)}\,\chi(\vartheta_1,\ \ldots,\ \vartheta_M)\, e^{i\varphi_0} + \\ + |\chi(\vartheta_1,\ \ldots,\ \vartheta_M)|^2], \tag{6.26}$$

где обозначено

$$h(\vartheta_1,\ \ldots,\ \vartheta_M) = |\Psi(\vartheta_1,\ \ldots,\ \vartheta_M)|^2 = \\ = \left|\frac{1}{2T}\int_{-T}^{T} a(t;\ \vartheta_1^0,\ \ldots,\ \vartheta_M^0)\,\overline{a(t;\ \vartheta_1,\ \ldots,\ \vartheta_M)}\,dt\right|^2; \tag{6.26′}$$

$$\chi(\vartheta_1,\ \ldots,\ \vartheta_M) = \frac{1}{2TA_0}\int_{-T}^{T} \overline{a(t;\ \vartheta_1,\ \ldots,\ \vartheta_M)}\,z_1(t)\,dt, \tag{6.26″}$$

причем функции h и Ψ — детерминированные, а $\chi(\vartheta_1, \ldots, \vartheta_M)$ — нормальная случайная величина с нулевым средним и дисперсией, равной обратной величине $d_T^2 = \frac{2TA_0^2}{N_0}$ отношения энергии сигнала к спектральной плотности шума.

При фиксированных значениях $\vartheta_1^0, \ldots, \vartheta_M^0$ функцию

$$\Psi(\vartheta_1,\ \ldots,\ \vartheta_M) = \frac{1}{2T}\int_{-T}^{T} a(t;\ \vartheta_1^0,\ \ldots,\ \vartheta_M^0)\,\overline{a(t;\ \vartheta_1,\ \ldots,\ \vartheta_M)}\,dt, \tag{6.27}$$

зависящую только от формы комплексной огибающей сигнала и ее параметров, называют *функцией неопределен-*

ности. Согласно неравенству Буняковского — Шварца (см. стр. 84 в первой книге)

$$|\Psi(\vartheta_1, \ldots, \vartheta_M)| \leqslant \left[\frac{1}{2T}\int_{-T}^{T}|a(t; \vartheta_1^0, \ldots, \vartheta_M^0|^2\,dt \times \right.$$

$$\left.\times\frac{1}{2T}\int_{-T}^{T}|a(t; \vartheta_1, \ldots, \vartheta_M)|^2\,dt\right]^{1/2}$$

или с учетом (6.21)

$$|\Psi(\vartheta_1, \ldots, \vartheta_M)| \leqslant 1. \qquad (6.27')$$

Максимальное значение, равное единице, функция $|\Psi|$ достигает при $\vartheta_i = \vartheta_i^0$, $i = 1, \ldots, M$ [см. (6.27)]. Это соответствует тому случаю, когда шумовое слагаемое в (6.25) исчезает, т. е. когда $d_T^2 \to \infty$. Таким образом,

$$\Psi(\vartheta_1^0, \ldots, \vartheta_M^0) = |\Psi(\vartheta_1^0, \ldots, \vartheta_M^0)| = 1. \qquad (6.27'')$$

Так как среднее по множеству реализаций функции Q равно

$$m_1\{Q(\vartheta_1, \ldots, \vartheta_M)\} = d_T^2\left[h(\vartheta_1, \ldots, \vartheta_M) + \frac{1}{d_T^2}\right]$$

и первый член в квадратных скобках в силу (6.27″) близок к единице при условии

$$d_T^2 = \frac{2TA_0^2}{N_0} \gg 1, \qquad (6.28)$$

то, ограничивая дальнейшие исследования условием (6.28), будем отыскивать максимум максиморум функции Q, пренебрегая в (6.26) последним членом. Вообще говоря, эта функция может иметь много максимумов в M-мерном пространстве неизвестных параметров ϑ_i. Если шумовое слагаемое в реализации отсутствует, то, как отмечалось выше, эта функция достигает абсолютного максимума при $\vartheta_i = \vartheta_i^0$, $i = 1, \ldots, M$. Наличие шума вызывает смещение этого максимума, но при выполнении условия (6.28) можно считать, что это смещение *невелико*, так что оценки $\hat{\vartheta}_i$, соответствующие абсолютному максимуму функции правдоподобия, *близки* к истинным величинам оцениваемых параметров ϑ_i. Учитывая это обстоятельство, разложим функцию $Q(\vartheta_1, \ldots, \vartheta_M)$ от M переменных в ряд Тейлора около точки $\vartheta_1^0, \ldots, \vartheta_M^0$. Ограничимся членами второго

порядка малости, опуская последнее слагаемое в (6.26) Далее имеем:

$$\frac{\partial h}{\partial \vartheta_k} = 0$$

в точке ($\vartheta_1^0, \ldots, \vartheta_M^0$), так как в этой точке функция $h(\vartheta_1, \ldots, \vartheta_M)$ имеет максимум. С учетом сделанных замечаний получим из (6.26)

$$Q(\vartheta_1, \ldots, \vartheta_M) = Q(\vartheta_1^0, \ldots, \vartheta_M^0) +$$

$$+ \sum_{k=1}^{M} (\vartheta_k - \vartheta_k^0) \left.\frac{\partial Q(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k}\right|_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0} +$$

$$+ \frac{1}{2} \sum_{k=1}^{M} \sum_{j=1}^{M} (\vartheta_k - \vartheta_k^0)(\vartheta_j - \vartheta_j^0) \left.\frac{\partial^2 Q(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k \, \partial \vartheta_j}\right|_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0},$$

где

$$Q(\vartheta_1^0, \ldots, \vartheta_M^0) = d_T^2 [1 + 2 \operatorname{Re} \chi(\vartheta_1^0, \ldots, \vartheta_M^0) e^{i\varphi_0}];$$

$$\left.\frac{\partial Q(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k}\right|_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0} = 2d_T^2 \operatorname{Re} e^{i\varphi_0} \times$$

$$\times \left[\frac{\partial \chi(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k} + \chi(\vartheta_1, \ldots, \vartheta_M) \times \right.$$

$$\left. \times \overline{\frac{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k}}\right]_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0}.$$

Дифференцируя полученное разложение функции Q по ϑ_k и приравнивая результат нулю, находим систему уравнений правдоподобия

$$\sum_{j=1}^{M} Q_{kj}(\vartheta_j - \vartheta_j^0) = \zeta_k, \quad k = 1, \ldots, M, \tag{6.29}$$

где обозначено

$$\zeta_k = -\left.\frac{\partial Q(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k}\right|_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0}, \tag{6.29'}$$

$$Q_{kj} = \left.\frac{\partial^2 Q(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k \, \partial \vartheta_j}\right|_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0}. \tag{6.29''}$$

Если детерминант матрицы **Q**, составленной из элементов Q_{kj}, отличен от нуля, то существует обратная матрица $\mathbf{Q}^{-1}$ с элементами, равными отношению алгебраических дополнений к детерминанту матрицы **Q**. Обозначив элементы обратной матрицы через q_{kj}, запишем решение

системы линейных уравнений (6.29) относительно ошибки $\hat{\vartheta}_j - \vartheta_j^0$ в виде

$$\hat{\vartheta}_j - \vartheta_j^0 = \sum_{i=1}^{M} \zeta_k q_{kj}, \quad j = 1, \ldots, M, \tag{6.30}$$

где в соответствии с принятым приближением [см. (6.26) и (6.28)]

$$\zeta_k = 2d_T^2 \operatorname{Re} \left\{ e^{i\varphi_0} \frac{1}{2TA_0} \int_{-T}^{T} \left[\overline{a(t; \vartheta_1^0, \ldots, \vartheta_M^0)} \times \right.\right.$$

$$\times \left(\frac{1}{2T} \int_{-T}^{T} a(t; \vartheta_1^0, \ldots, \vartheta_M^0) \frac{\overline{\partial a(t; \vartheta_1^0, \ldots, \vartheta_M^0)}}{\partial \vartheta_k} dt \right) +$$

$$\left. + \frac{\overline{\partial a(t; \vartheta_1^0, \ldots, \vartheta_M^0)}}{\partial \vartheta_k} \right] z_1(t)\, dt. \tag{6.30'}$$

Величины ζ_k и q_{kj} случайные, зависящие от шумового слагаемого реализации. Однако, сохраняя степень приближения, связанную с пренебрежением последним (квадратичным относительно χ) членом в (6.26), следует в выражении (6.29″) для Q_{kj} сохранить вклад лишь от детерминированного слагаемого (6.26). Тогда нужно положить при $d_T^2 \gg 1$

$$Q_{kj} = d_T^2 \frac{\partial^2 h(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k \, \partial \vartheta_j} \bigg|_{\vartheta_1 = \vartheta_1^0, \ldots, \vartheta_M = \vartheta_M^0}, \tag{6.31}$$

т. е. элементы матрицы $\mathbf{Q}$ равны произведению величины d_T^2 и смешанной второй производной квадрата модуля функции неопределенности в точке, соответствующей истинным значениям оцениваемых параметров.

Учитывая (6.31), приходим к выводу, что ошибки $\hat{\vartheta}_j - \vartheta_j^0$ [см. (6.30)] представляют взвешенные (с весами, равными элементам матрицы, обратной $\mathbf{Q}$) суммы *нормальных* случайных величин ζ_k с нулевыми средними и ковариациями

$$m_1\{\zeta_k \zeta_j\} = 2d_T^4 \operatorname{Re} m_1 \left\{ \left[\frac{\partial \chi(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_k} + \right.\right.$$

$$\left. + \chi(\vartheta_1, \ldots, \vartheta_M) \frac{\overline{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}}{\partial \vartheta_k} \right] \left[\frac{\overline{\partial \chi(\vartheta_1, \ldots, \vartheta_M)}}{\partial \vartheta_j} + \right.$$

$$\left.\left. + \overline{\chi(\vartheta_1, \ldots, \vartheta_M)} \frac{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}{\partial \vartheta_j} \right] \right\}_{\vartheta_1 = \vartheta_1^0, \ldots, \vartheta_M = \vartheta_M^0}. \tag{6.32}$$

При выводе формулы (6.32) учитываем, что для комплексной огибающей $m_1\{z_1(t)\,z_1(t')\}=0$ для любых t и t' [см. (5.37')].

Отсюда следует, во-первых, что оценки $\hat{\vartheta}_j$ в соответствии с общей теорией асимптотически (при $d_T^2 \gg 1$) *нормальны*. Средние значения этих оценок равны [см. (6.30)]

$$m_1\{\hat{\vartheta}_j\}=\vartheta_j^0, \tag{6.33}$$

т. е. оценки максимального правдоподобия (асимптотически при $d_T^2 \gg 1$) *несмещенные*. Далее, элементы корреляционной матрицы ошибок равны

$$m_1\{(\hat{\vartheta}_k-\vartheta_k^0)(\hat{\vartheta}_n-\vartheta_n^0)\}=\sum_{i=1}^{M}\sum_{j=1}^{M} q_{ki}q_{jn}m_1\{\zeta_i\zeta_j\}. \tag{6.34}$$

Найдем ковариацию $m_1\{\zeta_i\zeta_j\}$ из (6.32). Так как

$$m_1\left\{\overline{\chi(\vartheta_1^0,\ldots,\vartheta_M^0)}\frac{\partial\chi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_i}\right\}=\frac{1}{d_T^2}\frac{\partial\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_i}$$

и $m_1\{|\chi|^2\}=1/d_T^2$, то

$$m_1\{\zeta_i\zeta_j\}=2d_T^2\,\mathrm{Re}\left\{\left[\frac{\partial\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_j}\times\right.\right.$$

$$\times\overline{\frac{\partial\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_i}}+\frac{\partial\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_j}\frac{\partial\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_i}+$$

$$+\overline{\frac{\partial\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_j}}\,\overline{\frac{\partial\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_i}}+\frac{1}{2T}\times$$

$$\left.\left.\times\int_{-T}^{T}\frac{\partial a(t;\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_i}\overline{\frac{\partial a(t;\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_j}}\,dt\right]_{\vartheta_1=\vartheta_1^0,\ldots,\vartheta_M=\vartheta_M^0}\right\} \tag{6.35}$$

Кроме того, из (6.21) и (6.27) дифференцированием по ϑ_i и ϑ_j находим $\left(\text{следует иметь в виду, что } \left.\frac{\partial h}{\partial\vartheta_j}\right|_{\vartheta_1=\vartheta_1^0,\ldots,\vartheta_M=\vartheta_M^0}=0\right)$

$$\mathrm{Re}\left\{\left.\frac{\partial\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_j}\right|_{\vartheta_1=\vartheta_1^0,\ldots,\vartheta_M=\vartheta_M^0}\right\}=0, \tag{6.36}$$

$$\mathrm{Re}\left\{\left[\frac{\partial^2\Psi(\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_i\partial\vartheta_j}+\frac{1}{2T}\int_{-T}^{T}\frac{\partial a(t;\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_i}\times\right.\right.$$

$$\left.\left.\times\overline{\frac{\partial a(t;\vartheta_1,\ldots,\vartheta_M)}{\partial\vartheta_j}}\,dt\right]_{\vartheta_1=\vartheta_1^0,\ldots,\vartheta_M=\vartheta_M^0}\right\}=0, \tag{6.36'}$$

$$i,\ j=1,\ldots,M.$$

Из (6.26′) следует:

$$\frac{\partial^2 h(\vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_i \partial\vartheta_j} = 2\mathrm{Re}\left[\frac{\partial^2 \Psi(\vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_i \partial\vartheta_j} + \frac{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_i} \frac{\overline{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}}{\partial\vartheta_j}\right], \qquad (6.37)$$

откуда, учитывая (6.36′), получаем

$$\left.\frac{\partial^2 h(\vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_i \partial\vartheta_j}\right|_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0} = 2\,\mathrm{Re}\left\{\left[\frac{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_i} \frac{\overline{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}}{\partial\vartheta_j} - \frac{1}{2T}\int_{-T}^{T} \frac{\partial a(t; \vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_i} \times \frac{\overline{\partial a(t; \vartheta_1, \ldots, \vartheta_M)}}{\partial\vartheta_j}\,dt\right]_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0}\right\}.$$

Если использовать (6.36′), то (6.35) можно переписать в виде

$$m_1\{\zeta_i\zeta_j\} = 2d_T^2\,\mathrm{Re}\left\{\left[-\frac{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_j} \times \frac{\overline{\partial \Psi(\vartheta_1, \ldots, \vartheta_M)}}{\partial\vartheta_i} + \frac{1}{2T}\int_{-T}^{T} \frac{\partial a(t; \vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_i} \times \frac{\overline{\partial a(t; \vartheta_1, \ldots, \vartheta_M)}}{\partial\vartheta_j}\,dt\right]_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0}\right\}. \qquad (6.37'')$$

Сравнивая (6.37′) с (6.37″), получаем

$$m_1\{\zeta_i\zeta_j\} = -d_T^2 \left.\frac{\partial^2 h(\vartheta_1, \ldots, \vartheta_M)}{\partial\vartheta_i \partial\vartheta_j}\right|_{\vartheta_1=\vartheta_1^0, \ldots, \vartheta_M=\vartheta_M^0} \qquad (6.38)$$

или, учитывая (6.31), имеем

$$m_1\{\zeta_i\zeta_j\} = Q_{ij}. \qquad (6.38')$$

Подставляя (6.38′) в (6.34), находим

$$m_1\{(\hat{\vartheta}_k - \vartheta_k^0)(\hat{\vartheta}_n - \vartheta_n^0)\} = \sum_{i=1}^{M}\sum_{j=1}^{M} q_{ki}q_{jn}Q_{ij}. \qquad (6.39)$$

Но так как q_{ij} — элементы матрицы $\mathbf{Q}^{-1}$, обратной $\mathbf{Q}$, то из элементарного матричного равенства $\mathbf{Q}^{-1}\mathbf{Q} = \mathbf{I}$ следует:

$$\sum_{j=1}^{M} q_{jn} Q_{ij} \quad \begin{cases} 1, & i = n, \\ 0, & i \neq n. \end{cases} \tag{6.39'}$$

Соединяя (6.39) и (6.39′), приходим к окончательной формуле, определяющей любой из элементов корреляционной матрицы ошибок:

$$m_1 \{(\hat{\vartheta}_k - \vartheta_k^0)(\hat{\vartheta}_n - \vartheta_n^0)\} = -q_{kn}. \tag{6.40}$$

Итак, с точностью до знака элементы корреляционной матрицы ошибок совпадают с элементами матрицы, обратной матрице, составленной из смешанных вторых производных логарифма отношения правдоподобия (которые лишь множителем d_T^2 отличаются от смешанных производных квадрата модуля функций неопределенности) в точке, соответствующей истинным значениям $\vartheta_1^0, \ldots, \vartheta_M^0$ оцениваемых параметров. Отсюда следует, что корреляционная матрица ошибок совпадает с матрицей, обратной информационной матрице Фишера [см. (2.98)], т. е. оценки (6.30) *совместно эффективные* (при $d_T^2 \gg 1$).

6.1.4. Измерение времени прихода сигнала. В качестве первого примера, иллюстрирующего методику § 6.1.3 получения асимптотических (при $d_T^2 \to \infty$) оценок максимального правдоподобия, рассмотрим узкополосный сигнал вида *)

$$s(t;\ A,\ \varphi,\ \tau) = A\,\mathrm{Re}\, a(t-\tau)\, e^{-i\varphi} e^{i\omega_0 t}, \tag{6.41}$$

в котором параметр τ может быть (следуя радиолокационной терминологии) назван *временем прихода* сигнала.

Функция неопределенности такого сигнала согласно (6.27) равна

$$\Psi(\tau - \tau_0) = \frac{1}{2T} \int_{-T}^{T} a(t-\tau_0)\, \overline{a(t-\tau)}\, dt, \tag{6.42}$$

где τ_0 — истинное значение параметра τ. Переписывая (6.42) в виде

$$\Psi(\tau - \tau_0) = \frac{1}{2T} \int_{-T-\tau_0}^{T-\tau_0} a(t)\, \overline{a(t-\tau+\tau_0)}\, dt,$$

---

*) $a(t) = u(t)\, e^{i\Phi(t)}$, где $u(t)$ и $\Phi(t)$ — действительные функции, представляющие амплитудную и фазовую модуляцию сигнала.

замечаем, что функция неопределенности в этом случае совпадает с временнóй корреляционной функцией комплексной огибающей сигнала (см. § 4.2 в первой книге). Матрица **Q** теперь состоит из одного элемента, равного [см. (6.31) и (6.36)]

$$\frac{d^2}{d\tau^2}\,|\,\Psi(\tau-\tau_0)\,|^2_{\tau=\tau_0}=2\,\mathrm{Re}\,\{\Psi''(0)-[\Psi'(0)]^2\}. \qquad (6.43)$$

Введем энергетический спектр огибающей сигнала (см. (4.86) в первой книге)

$$F_a(\omega)=\frac{1}{T}\left|\int\limits_{-T}^{T} a(t)\,\mathrm{e}^{-i\omega t}dt\right|^2. \qquad (6.44)$$

Тогда

$$\Psi(\tau-\tau_0)=\frac{1}{4\pi}\int\limits_{-\infty}^{\infty}F_a(\omega)\,\mathrm{e}^{i\omega(\tau-\tau_0)}\,d\omega, \qquad (6.45)$$

откуда находим

$$-2i\Psi'(0)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\omega F_a(\omega)\,d\omega=\omega^*, \qquad (6.46)$$

$$-2\Psi''(0)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\omega^2F_a(\omega)\,d\omega=\omega_*^2, \qquad (6.47)$$

где ω^* и ω_*^2 — средняя частота и второй момент энергетического спектра огибающей сигнала с учетом нормировки (6.21) (обозначения соответствуют § 8.6.1 первой книги).

Подставляя (6.46) и (6.47) в (6.43), получаем

$$\left[-\frac{d^2}{d\tau^2}\,|\,\Psi(\tau-\tau_0)\,|^2\right]_{\tau=\tau_0}=\omega_*^2-\omega^{*2}=\beta^2, \qquad (6.48)$$

где β^2 — среднеквадратическая девиация (отклонение от среднего) частоты спектра огибающей. Величину β иногда принимают в качестве количественной меры полосы спектра огибающей.

Так как в рассматриваемом случае обратная матрица $\mathbf{Q}^{-1}$ содержит только один элемент, равный $-\frac{1}{d_T^2\beta^2}$, то из (6.30), (6.30′) и (6.46) находим следующее асимптотиче-

ское выражение (при $d_T^2 \gg 1$) оценки $\hat{\tau}$ максимального правдоподобия времени прихода сигнала:

$$\hat{\tau} - \tau_0 =$$

$$= \frac{2}{\beta^2} \operatorname{Re} \left\{ e^{i\varphi_0} \frac{1}{2TA_0} \int_{-T}^{T} \left[\overline{\frac{\partial a(t-\tau)}{\partial \tau}} \bigg|_{\tau=\tau_0} - \right. \right.$$

$$\left. \left. - i\omega^* \overline{a(t-\tau_0)} \right] z(t)\, dt \right\}. \qquad (6.49)$$

Эта оценка несмещенная, а ее дисперсия согласно (6.40) равна

$$M_2\{\hat{\tau}\} = \frac{1}{d_T^2 \beta^2}. \qquad (6.50)$$

При симметричном спектре огибающей $\omega^* = 0$ выражение (6.49) можно переписать в виде

$$\hat{\tau} - \tau_0 = \frac{2}{\omega_*^2} \operatorname{Re} \left[e^{i\varphi_0} \frac{1}{2TA_0} \int_{-T}^{T} \overline{\frac{\partial a(t-\tau)}{\partial \tau}} \bigg|_{\tau=\tau_0} z(t)\, dt \right]. \qquad (6.49')$$

Из (6.23′) и (6.23″) находим также оценки амплитуды и фазы:

$$\hat{A} = \left| \frac{1}{2T} \int_{-T}^{T} a(t-\hat{\tau})\, z(t)\, dt \right|, \qquad (6.51)$$

$$\hat{\varphi} = \operatorname{arctg} \frac{\operatorname{Im} \int_{-T}^{T} a(t-\hat{\tau})\, z(t)\, dt}{\operatorname{Re} \int_{-T}^{T} a(t-\hat{\tau})\, z(t)\, dt}, \qquad (6.51')$$

в которых $\hat{\tau}$ подставляется из (6.49).

6.1.5. Совместное измерение времени прихода и допплеровского смещения частот. В качестве второй иллюстрации материала § 6.1.3 рассмотрим узкополосный сигнал более сложного вида, чем (6.41):

$$s(t;\ A,\ \varphi,\ \tau,\ \Omega) = A \operatorname{Re} a(t-\tau)\, e^{-i\varphi}\, e^{i\Omega(t-\tau)}\, e^{i\omega_0(t-\tau)}. \qquad (6.52)$$

Здесь, как и в (6.41), параметр τ — время прихода сигнала, а

$$\Omega = \omega_1 - \omega_0$$

— смещение частоты, которое назовем *допплеровским*. Задача состоит в совместной оценке параметров τ и Ω.

Функция неопределенности *) сигнала (6.52) равна в соответствии с (6.27)

$$\Psi(\tau-\tau_0,\ \Omega-\Omega_0)=\frac{1}{2T}\int_{-T}^{T} a(t-\tau_0)\overline{a(t-\tau)}\times$$
$$\times e^{-i\Omega_0(t-\tau_0)}e^{i\Omega(t-\tau)}\,dt=e^{i\Omega(\tau-\tau_0)}\times$$
$$\times\frac{1}{2T}\int_{-T-\tau_0}^{T-\tau_0} a(t)\overline{a(t-\tau+\tau_0)}\,e^{i(\Omega-\Omega_0)t}\,dt. \qquad (6.53)$$

Вводя спектр огибающей сигнала

$$f_T(\omega)=\int_{-T}^{T} a(t)\,e^{-i\omega t}\,dt,$$

можно функцию неопределенности переписать в виде

$$\Psi(\tau-\tau_0,\ \Omega-\Omega_0)=$$
$$=e^{i\Omega(\tau-\tau_0)}\frac{1}{4\pi T}\int_{-\infty}^{\infty} f_T(\omega)\overline{f_T(\omega-\Omega+\Omega_0)}\,e^{i\omega(\tau-\tau_0)}\,d\omega. \qquad (6.54)$$

Элементы матрицы Q размером 2×2 равны:

$$Q_{11}=d_T^2\frac{\partial^2}{\partial\tau^2}|\Psi(\tau-\tau_0,\ \Omega-\Omega_0)|^2_{\tau=\tau_0,\ \Omega=\Omega_0}, \qquad (6.55)$$

$$Q_{12}=Q_{21}=d_T^2\frac{\partial^2}{\partial\tau\,\partial\Omega}|\Psi(\tau-\tau_0,\ \Omega-\Omega_0)|^2_{\tau=\tau_0,\ \Omega=\Omega_0}, \qquad (6.55')$$

$$Q_{22}=d_T^2\frac{\partial^2}{\partial\Omega^2}|\Psi(\tau-\tau_0,\ \Omega-\Omega_0)|^2_{\tau=\tau_0,\ \Omega=\Omega_0}. \qquad (6.55'')$$

Из (6.54) и (6.48) следует, что

$$Q_{11}=d_T^2\frac{\partial^2}{\partial\tau^2}|\Psi(\tau-\tau_0,\ 0)|^2_{\tau=\tau_0}=-d_T^2\beta^2. \qquad (6.56)$$

Величину Q_{22} находим из (6.53):

$$Q_{22}=d_T^2\frac{\partial^2}{\partial\Omega^2}|\Psi(0,\ \Omega-\Omega_0)|^2_{\Omega=\Omega_0}=$$
$$=d_T^2\left[\left(\frac{1}{2T}\int_{-T}^{T} t\,|a(t-\tau_0)|^2\,dt\right)^2-\right.$$
$$\left.-\frac{1}{2T}\int_{-T}^{T} t^2\,|a(t-\tau_0)|^2\,dt\right]. \qquad (6.57)$$

---

*) Функция $\Psi(\tau,\ \Omega)$ двух переменных (времени и частоты) была впервые введена Вудвордом как функция неопределенности [2].

Вводя по аналогии с ω^* и ω_*^2 *среднее время* t^* и *второй момент* t_*^2 сигнала и обозначая среднеквадратическую длительность сигнала

$$\alpha^2 = t_*^2 - (t^*)^2 = \frac{1}{2T} \int_{-T}^{T} t^2 |a(t-\tau_0)|^2 \, dt - \left(\frac{1}{2T} \int_{-T}^{T} t \, |a(t-\tau_0)|^2 \, dt \right)^2, \tag{6.58}$$

перепишем (6.57) в виде

$$Q_{22} = -d_T^2 \alpha^2.$$

Из (6.53) и (6.54) находим также

$$Q_{12} = d_T^2 \left\{ -\frac{1}{2\pi} \int_{-\infty}^{\infty} \omega F_s(\omega) \, d\omega \cdot \frac{1}{2T} \int_{-T}^{T} t \, |a(t-\tau_0)|^2 \, dt + \frac{i}{2T} \int_{-T}^{T} [a(t-\tau_0) \overline{a'(t-\tau_0)} - \overline{a(t-\tau_0)} \, a'(t-\tau_0)] \, t \, dt \right\}$$

или

$$Q_{12} = -d_T^2 (\omega^* t^* - \lambda^*), \tag{6.59}$$

где через λ^* обозначена безразмерная величина, характеризующая *частотную модуляцию* сигнала *),

$$\lambda^* = \operatorname{Im} \frac{1}{T} \int_{-T}^{T} t a(t-\tau_0) \overline{a'(t-\tau_0)} \, dt. \tag{6.60}$$

Таким образом, матрица $\mathbf{Q}$ имеет следующий вид:

$$\mathbf{Q} = d_T^4 \begin{pmatrix} -\beta^2 & -\omega^* t^* + \lambda^* \\ -\omega^* t^* + \lambda^* & -\alpha^2 \end{pmatrix}. \tag{6.61}$$

---

*) Действительно, если $a(t) = u(t) \, e^{i\Phi(t)}$, то $\lambda^* = \frac{1}{T} \int_{-T}^{T} t u^2(t) \times \frac{d\Phi(t)}{dt} \, dt$.

Нетрудно найти обратную матрицу. Вычисляя детерминант матрицы

$$\det \mathbf{Q} = d_T^4 [\alpha^2\beta^2 - (\omega^* t^* - \lambda^*)^2], \tag{6.62}$$

находим корреляционную матрицу ошибок измерения времени прихода и допплерового сдвига частоты

$$\mathbf{Q}^{-1} = \frac{1}{d_T^4 [\alpha^2\beta^2 - (\omega^* t^* - \lambda^*)^2]^2} \begin{pmatrix} -\alpha^2 & \omega^* t^* - \lambda^* \\ \omega^* t^* - \lambda^* & -\beta^2 \end{pmatrix}. \tag{6.63}$$

На главной диагонали матрицы $\mathbf{Q}^{-1}$ находятся дисперсии оценок времени прихода и допплерова сдвига частоты, а на побочной — ковариация этих оценок (с точностью до знака; см. (6.40)).

Можно несколько упростить формулу (6.63), если надлежащим выбором начала координат в плоскости время — частота обратить в нуль величины t^* и ω^*. Тогда

$$\mathbf{Q}^{-1} = \frac{1}{d_T^4 (\omega_*^2 t_*^2 - \lambda^{*2})^2} \begin{pmatrix} -t_*^2 & -\lambda^* \\ -\lambda^* & -\omega_*^2 \end{pmatrix}. \tag{6.64}$$

Используя (6.30), (6.30′), (6.64), находим следующее асимптотическое выражение (при $d_T^2 \gg 1$) совместных оценок времени прихода сигнала и допплерова сдвига частоты:

$$\hat{\tau} - \tau_0 = \frac{2}{\omega_*^2 t_*^2 - \lambda^{*2}} \operatorname{Re} \left\{ e^{i\varphi_0} \frac{1}{2TA_0} \times \right.$$

$$\left. \times \int_{-T}^{T} \left[t_*^2 \left. \frac{\partial a(t-\tau)}{\partial \tau} \right|_{\tau=\tau_0} - i\lambda^* t\, \overline{a(t-\tau_0)} \right] e^{-i\Omega_0(t-\tau_0)} z(t)\, dt \right\}, \tag{6.65}$$

$$\hat{\Omega} - \Omega_0 = \frac{2}{\omega_*^2 t_*^2 - \lambda^{*2}} \operatorname{Re} \left\{ e^{i\varphi_0} \frac{1}{2TA_0} \times \right.$$

$$\times \int_{-T}^{T} \left[\lambda^* \left. \frac{\partial a(t-\tau)}{\partial \tau} \right|_{\tau=\tau_0} - i\omega_*^2 t\, \overline{a(t-\tau_0)} \right] \times$$

$$\left. \times e^{-i\Omega_0(t-\tau_0)} z(t)\, dt \right\}. \tag{6.65′}$$

В том случае, когда частотная модуляция сигнала отсутствует ($\lambda^* = 0$), оценка времени прихода и оценка допплеровского сдвига частоты становятся некоррелированными, а их дисперсии равны [см. (6.40), а также (6.48)]

$$M_2\{\hat{\tau}\} = \frac{1}{d_T^2 \omega_*^2}, \qquad M_2\{\hat{\Omega}\} = \frac{1}{d_T^2 t_*^2}. \tag{6.65''}$$

6.2. БАЙЕСОВСКИЕ ОЦЕНКИ СЛУЧАЙНЫХ ПАРАМЕТРОВ СИГНАЛА

6.2.1. Совместные оценки амплитуды и фазы. Переходим к совместным оценкам случайных параметров квазидетерминированного сигнала по наблюдаемой реализации аддитивной смеси этого сигнала с нормальным стационарным шумом, среднее значение которого равно нулю, а корреляционная функция равна $B(\tau)$.

Из общей теории следует, что при квадратичной функции потерь (а иногда и при более общих условиях; см. § 3.6.4) байесовские оценки совокупности зависимых параметров случайного процесса представляют условные средние этих параметров, когда наблюдается реализация $x(t)$ процесса на некотором интервале времени [см. (3.186)].

Байесовская оценка случайной амплитуды a квазидетерминированного сигнала $as(t)$ в аддитивном нормальном шуме была подробно рассмотрена в § 3.6.3. Там было показано, что при нормальном распределении параметра указанная оценка представляет среднее взвешенное двух величин: априорного среднего и оценки максимального правдоподобия (исследованной в § 3.6.2).

Предположим теперь, что сигнал представляет узкополосный квазидетерминированный случайный процесс $s(t) = a\cos(\omega_0 t - \varphi)$ известной частоты ω_0, амплитуда a и фаза φ которого случайны с известным априорным распределением $w_2(a, \varphi)$. Нормальный шум также узкополосный. Воспользовавшись комплексными представлениями реализации $x(t)$ и сигнала $s(t)$

$$x(t) = \operatorname{Re} z(t)\, e^{i\omega_0 t}, \quad s(t;\ a,\ \varphi) = \operatorname{Re} a\, e^{-i\varphi}\, e^{i\omega_0 t}, \tag{6.66}$$

запишем выражение функционала правдоподобия в виде [см. (3.78) и (5.54)]

$$l[z(t)\,|\,a,\varphi]=\exp\left[\operatorname{Re}\int_{-T}^{T}V(t;\ a,\ \varphi)\,\overline{z(t)}\,e^{-i\omega_0 t}\,dt\right]\times$$

$$\times\exp\left[-\frac{a}{2}\operatorname{Re}\int_{-T}^{T}V(t;\ a,\ \varphi)\,e^{-i\omega_0 t}\,e^{i\varphi}\,dt\right],\qquad(6.67)$$

где $V(t;\ a,\ \varphi)$ — решение интегрального уравнения

$$\int_{-T}^{T}B_z(t-u)\,V(u;\ a,\ \varphi)\,du=a\,e^{-i\varphi}\,e^{i\omega_0 t},\qquad(6.67')$$

$$|t|\leqslant T.$$

Вводя функцию $U(t)=V(t;\ a,\ \varphi)\,ae^{i\varphi}$, перепишем (6.67) и (6.67′) следующим образом:

$$l[z(t)\,|\,a,\varphi]=\exp\left[a\operatorname{Re}e^{-i\varphi}\int_{-T}^{T}U(t)\,\overline{z(t)}\,e^{-i\omega_0 t}\,dt\right]\times$$

$$\times\exp\left[-\frac{a^2}{2}\operatorname{Re}\int_{-T}^{T}U(t)\,e^{-i\omega_0 t}\,dt\right]=$$

$$=\exp\left[a\left|\int_{-T}^{T}U(t)\,\overline{z(t)}\,e^{-i\omega_0 t}\,dt\right|\cos(\varphi-\psi_T)\right]\times$$

$$\times\exp\left[-\frac{a^2}{2}\operatorname{Re}\int_{-T}^{T}U(t)\,e^{-i\omega_0 t}\,dt\right],\qquad(6.68)$$

где

$$\psi_T=\operatorname{arctg}\frac{\operatorname{Im}\int_{-T}^{T}U(t)\,\overline{z(t)}\,e^{-i\omega_0 t}\,dt}{\operatorname{Re}\int_{-T}^{T}U(t)\,\overline{z(t)}\,e^{-i\omega_0 t}\,dt},\qquad(6.68')$$

а $U(t)$ — решение интегрального уравнения

$$\int_{-T}^{T}B_z(t-y)\,U(y)\,dy=e^{i\omega_0 t},\quad |t|\leqslant T,\qquad(6.69)$$

где $B_z(\tau)$ — комплексная огибающая корреляционной функции шума [см. (5.35′)].

Введем по аналогии с § 5.2.2 обозначения:

$$r_T^2 = \sigma_a^2 \int_{-T}^{T} \int_{-T}^{T} U(t)\,\overline{U(y)}\,\overline{z(t)}\,z(y)\, e^{-i\omega_0(t-y)}\, dt\, dy, \qquad (6.70)$$

$$d_T^2 = \sigma_a^2 \int_{-T}^{T} \overline{U(t)}\, e^{i\omega_0 t}\, dt =$$

$$= \sigma_a^2 \int_{-T}^{T} \int_{-T}^{T} B_z(t-y)\, U(y)\, \overline{U(t)}\, dy\, dt > 0, \qquad (6.71)$$

где σ_a^2 — средний квадрат амплитуды.

Используя (6.68) и введенные обозначения, находим байесовские оценки амплитуды и фазы узкополосного сигнала в аддитивном нормальном узкополосном шуме:

$$\hat{a} = m_1\{a \mid z(t)\} =$$

$$= \frac{\int_{-\infty}^{\infty} \int_{-\pi}^{\pi} a w_2(a, \varphi)\, l[z(t) \mid a, \varphi]\, d\varphi\, da}{\int_{-\infty}^{\infty} \int_{-\pi}^{\pi} w_2(a, \varphi)\, l[z(t) \mid a, \varphi]\, d\varphi\, da} =$$

$$= \frac{\int_{-\infty}^{\infty} \int_{-\pi}^{\pi} a w_2(a, \varphi)\, e^{-\frac{a^2 d_T^2}{2\sigma_a^2}} \exp\left[\frac{a r_T}{\sigma_a} \cos(\varphi - \psi_T)\right] d\varphi\, da}{\int_{-\infty}^{\infty} \int_{-\pi}^{\pi} w_2(a, \varphi)\, e^{-\frac{a^2 d_T^2}{2\sigma_a^2}} \exp\left[\frac{a r_T}{\sigma_a} \cos(\varphi - \psi_T)\right] d\varphi\, da}, \qquad (6.72)$$

$$\hat{\varphi} = m_1\{\varphi \mid z(t)\} =$$

$$= \frac{\int_{-\infty}^{\infty} \int_{-\pi}^{\pi} \varphi w_2(a, \varphi)\, e^{-\frac{a^2 d_T^2}{2\sigma_a^2}} \exp\left[\frac{a r_T}{\sigma_a} \cos(\varphi - \psi_T)\right] d\varphi\, da}{\int_{-\infty}^{\infty} \int_{-\pi}^{\pi} w_2(a, \varphi)\, e^{-\frac{a^2 d_T^2}{2\sigma_a^2}} \exp\left[\frac{a r_T}{\sigma_a} \cos(\varphi - \psi_T)\right] d\varphi\, da}. \qquad (6.73)$$

Если аддитивный шум белый с интенсивностью N_0, то решение интегрального уравнения (6.69) равно $U(t) = \frac{1}{N_0} e^{i\omega_0 t}$ и тогда величины r_T и d_T в приведенных формулах равны [см. (6.70) и (6.71)]

$$r_T = \frac{\sigma_a}{N_0} \left| \int_{-T}^{T} z(t)\, dt \right| =$$

$$= \frac{\sigma_a}{N_0} \left[\left(\int_{-T}^{T} A(t)\, dt \right)^2 + \left(\int_{-T}^{T} C(t)\, dt \right)^2 \right]^{\frac{1}{2}}, \qquad (6.74)$$

где $A(t)$ и $C(t)$ — квадратурные составляющие наблюдаемой реализации (см. § 5.2.3);

$$d_T^2 = \frac{2T\sigma_a^2}{N_0}. \qquad (6.74')$$

Если фаза распределена равномерно и независима от амплитуды, то

$$\hat{a} = \sigma_a \frac{\int_{-\infty}^{\infty} x w_1(\sigma_a x) I_0(r_T x) e^{-\frac{d_T^2 x^2}{2}} dx}{\int_{-\infty}^{\infty} w_1(\sigma_a x) I_0(r_T x) e^{-\frac{d_T^2 x^2}{2}} dx}, \qquad (6.75)$$

$$\hat{\varphi} = \frac{\frac{1}{2\pi} \int_{-\infty}^{\infty} w_1(\sigma_a x) e^{-\frac{x^2 d_T^2}{2}} \left(\int_{-\pi}^{\pi} \varphi \exp\left[r_T x \cos(\varphi - \psi_T)\right] d\varphi \right) dx}{\int_{-\infty}^{\infty} w_1(\sigma_a x) I_0(r_T x) e^{-\frac{x^2 d_T^2}{2}} dx}, \qquad (6.75')$$

где $w_1(x)$ — функция распределения амплитуды.

Для сильного сигнала байесовская оценка (6.75) амплитуды сигнала может быть преобразована путем использования асимптотической формулы для бесселевой функции

(см., например, стр. 123 в первой книге) к виду

$$\hat{a} \sim \sigma_a \frac{\int\limits_{-\infty}^{\infty} \sqrt{x}\, w_1(\sigma_a x) \exp\left(-\frac{d_T^2 x^2}{2} + r_T x\right) dx}{\int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{x}}\, w_1(\sigma_a x) \exp\left(-\frac{d_T^2 x^2}{2} + r_T x\right) dx} =$$

$$= \sigma_a \frac{\int\limits_{-\infty}^{\infty} \sqrt{x}\, w_1(\sigma_a x) \exp\left[-\frac{d_T^2}{2}\left(x - \frac{r_T}{d_T^2}\right)^2\right] dx}{\int\limits_{-\infty}^{\infty} \frac{1}{\sqrt{x}}\, w_1(\sigma_a x) \exp\left[-\frac{d_T^2}{2}\left(x - \frac{r_T}{d_T^2}\right)^2\right] dx}.$$

При $d_T \gg 1$ функция $\frac{d_T}{\sqrt{2\pi}} \exp\left[-\frac{d_T^2}{2}\left(x - \frac{r_T}{d_T^2}\right)^2\right]$ близка к дельта-функции. Поэтому при условии, что плотность $w_1(\sigma_a x)$ непрерывна в окрестности точки $x = \frac{r_T}{d_T^2}$ и не равна нулю в этой точке, получаем, используя фильтрующее свойство дельта-функции (см. приложение III в первой книге), следующую асимптотическую формулу:

$$\hat{a} \sim \frac{\sigma_a r_T}{d_T^2} = \frac{\left| \int\limits_{-T}^{T} U(t)\, \overline{z(t)}\, e^{-i\omega_0 t}\, dt \right|}{\int\limits_{-T}^{T} \int\limits_{-T}^{T} B_z(t-y)\, U(y)\, \overline{U(t)}\, dy\, dt}. \qquad (6.76)$$

Аналогично получаем асимптотическую формулу для оценки фазы при $d_T^2 \gg 1$:

$$\hat{\varphi} \sim \psi_T = \operatorname{arctg} \frac{\operatorname{Im} \int\limits_{-T}^{T} \overline{U(t)}\, z(t)\, e^{i\omega_0 t}\, dt}{\operatorname{Re} \int\limits_{-T}^{T} \overline{U(t)}\, z(t)\, e^{i\omega_0 t}\, dt}. \qquad (6.76')$$

Таким образом, байесовские оценки амплитуды и фазы совпадают с оценками максимального правдоподобия этих

параметров [см. (6.5) и (6.5′)]

$$\hat{a} \sim \left| \frac{1}{2T} \int_{-T}^{T} z(t)\, dt \right|, \tag{6.76″}$$

$$\hat{\varphi} \sim \operatorname{arctg} \frac{\operatorname{Im} \int_{-T}^{T} z(t)\, dt}{\operatorname{Re} \int_{-T}^{T} z(t)\, dt}, \tag{6.76‴}$$

как и в формулах (6.9) и (6.7).

6.2.2. Совместные оценки конечного числа параметров квазидетерминированного сигнала. Рассмотрим задачу об определении совместных байесовских оценок параметров квазидетерминированного сигнала в более общей, чем в предыдущем разделе, форме (см. § 3.6.6). Пусть сигнал заданной формы зависит от m случайных параметров, и предположим, что он может быть представлен в виде

$$s(t;\ \vartheta_1, \ldots, \vartheta_m) = \sum_{j=1}^{m} \vartheta_j s_j(t), \tag{6.77}$$

где $s_j(t)$ — известные функции, а ϑ_j — случайные параметры, совместное распределение которых задано распределением $w_m(\vartheta_1, \ldots, \vartheta_m)$. Наблюдается на интервале $(-T, T)$ реализация $x(t)$ суммы $s(t;\ \vartheta_1, \ldots, \vartheta_m)$ и нормального случайного процесса с нулевым средним и известной корреляционной функцией $B(u, v)$. Необходимо найти байесовские оценки параметров $\vartheta_1, \ldots, \vartheta_m$.

Логарифм функционала отношения правдоподобия при фиксированных $\vartheta_1, \ldots, \vartheta_m$ в рассматриваемом случае в соответствии с (3.157) равен

$$\ln l\,[x(t) \mid \vartheta_1, \ldots, \vartheta_m] = \int_{-T}^{T} V(t;\ \vartheta_1, \ldots, \vartheta_m) \left[x(t) - \frac{1}{2} s(t;\ \vartheta_1, \ldots, \vartheta_m) \right] dt, \tag{6.78}$$

где $V(t;\ \vartheta_1, \ldots, \vartheta_m)$ — решение интегрального уравнения

$$\int_{-T}^{T} B(t, u) V(u;\ \vartheta_1, \ldots, \vartheta_m)\, du = s(t;\ \vartheta_1, \ldots, \vartheta_m), \quad |t| \leqslant T. \tag{6.78′}$$

Подставляя (6.77) в (6.78′), заменим (6.78′) системой уравнений

$$\int_{-T}^{T} B(t, u) V_i(u)\, du = s_i(t), \qquad |t| \leqslant T, \qquad i = 1, \ldots, m, \tag{6.79}$$

причем

$$V(t;\ \vartheta_1, \ldots, \vartheta_m) = \sum_{j=1}^{m} \vartheta_j V_j(t). \tag{6.80}$$

Выражение (6.78) может быть теперь переписано в виде

$$\ln l\,[x(t)\,|\,\vartheta_1, \ldots, \vartheta_m] = \sum_{j=1}^{m} \vartheta_j \int_{-T}^{T} x(t) \times$$

$$\times V_j(t)\, dt - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \vartheta_i \vartheta_j \int_{-T}^{T} V_i(t)\, s_j(t)\, dt$$

или в обозначениях (3.206), (3.206′)

$$\ln l\,[x(t)\,|\,\vartheta_1, \ldots, \vartheta_m] = \sum_{j=1}^{m} \vartheta_j x_{Tj} -$$

$$- \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \vartheta_i \vartheta_j s_{Tij}. \tag{6.81}$$

Теперь, используя (3.186), можно записать общее выражение байесовских оценок параметров при квадратичной функции потерь:

$$\hat{\vartheta}_k = m_1\{\vartheta_k\,|\,x(t)\} =$$

$$= \Bigg[\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \vartheta_k w_m(\vartheta_1, \ldots, \vartheta_m) \exp\Bigg(\sum_{j=1}^{m} \vartheta_j x_{Tj} -$$

$$- \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \vartheta_i \vartheta_j s_{Tij} \Bigg) d\vartheta_1 \ldots d\vartheta_m \Bigg] \times$$

$$\times \Bigg[\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} w_m(\vartheta_1, \ldots, \vartheta_m) \exp\Bigg(\sum_{j=1}^{m} \vartheta_j x_{Tj} - \tag{6.82}$$

$$- \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \vartheta_i \vartheta_j s_{Tij} \Bigg) d\vartheta_1 \ldots d\vartheta_m \Bigg]^{-1},$$

$$k = 1, \ldots, m,$$

или в векторной форме

$$\hat{\boldsymbol{\vartheta}} = \frac{\int\limits_{G_\vartheta} \boldsymbol{\vartheta} w(\boldsymbol{\vartheta}) \exp\left(\boldsymbol{\vartheta}'\mathbf{x}_T - \frac{1}{2}\boldsymbol{\vartheta}'\mathbf{s}_T\boldsymbol{\vartheta}\right) d\boldsymbol{\vartheta}}{\int\limits_{G_\vartheta} w(\boldsymbol{\vartheta}) \exp\left(\boldsymbol{\vartheta}'\mathbf{x}_T - \frac{1}{2}\boldsymbol{\vartheta}'\mathbf{s}_T\boldsymbol{\vartheta}\right) d\boldsymbol{\vartheta}}. \qquad (6.82')$$

Если аддитивный шум белый, то $V_i(t) = \frac{1}{N_0} s_i(t)$ и

$$x_{Tj} = \frac{1}{N_0} \int\limits_{-T}^{T} x(t) s_j(t)\, dt, \qquad (6.83)$$

$$s_{Tij} = \frac{1}{N_0} \int\limits_{-T}^{T} s_i(t) s_j(t)\, dt. \qquad (6.84)$$

В том случае, когда функции $s_j(t)$, $j = 1, \ldots, m$, в (6.77) ортогональны на интервале $(-T, T)$,

$$s_{Tij} = \begin{cases} \frac{E_i}{N_0}, & i = j, \\ 0, & i \neq j, \end{cases} \qquad (6.85)$$

где E_i — энергия сигнала $s_i(t)$ на интервале наблюдения,

$$E_i = \int\limits_{-T}^{T} s_i^2(t)\, dt. \qquad (6.86)$$

Для этого частного случая формула (6.82) значительно упрощается:

$$\begin{aligned} \hat{\vartheta}_k = \Bigg[\int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} \vartheta_k w_m(\vartheta_1, \ldots, \vartheta_m) \times \\ \times \exp\left[-\frac{1}{2N_0} \sum_{j=1}^{m} E_j \left(\vartheta_j - \frac{N_0 x_{Tj}}{E_j}\right)^2\right] d\vartheta_1 \ldots d\vartheta_m \Bigg] \times \\ \times \Bigg[\int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} w_m(\vartheta_1, \ldots, \vartheta_m) \times \\ \times \exp\left[-\frac{1}{2N_0} \sum_{j=1}^{m} E_j \left(\vartheta_j - \frac{N_0 x_{Tj}}{E_j}\right)^2\right] d\vartheta_1 \ldots d\vartheta_m \Bigg]^{-1} \end{aligned} \qquad (6.87)$$

При $\frac{E_j}{N_0} \to \infty$, $j=1, \ldots, m$ (сильный сигнал),

$$\frac{E_1 \ldots E_m}{(2\pi N_0)^{m/2}} \exp\left[-\frac{1}{2N_0}\sum_{j=1}^{m} E_j\left(\vartheta_j - \frac{N_0 x_{Tj}}{E_j}\right)^2\right] \sim$$

$$\sim \prod_{j=1}^{m} \delta\left(\vartheta_j - \frac{N_0 x_{Tj}}{E_j}\right).$$

Тогда, используя фильтрующее свойство дельта-функции, из (6.87) находим при $E_j/N_0 \to \infty$ асимптотические байесовские оценки параметров сильного сигнала:

$$\hat{\vartheta}_k \sim \frac{x_{Tk}}{E_k/N_0} = \frac{\int_{-T}^{T} x(t)\, s_k(t)\, dt}{\int_{-T}^{T} s_k^2(t)\, dt}, \qquad k=1, \ldots, m. \qquad (6.88)$$

Полученная оценка $\hat{\vartheta}_k$ совпадает с оценкой максимального правдоподобия параметра ϑ_k сигнала $\vartheta_k s_k(t)$ [см. (3.168)].

Можно показать [13], что и в общем случае из (6.82) при сильном сигнале имеет место формула *)

$$\hat{\boldsymbol{\vartheta}} = \mathbf{x}_T \mathbf{S}_T^{-1}, \qquad (6.89)$$

т. е. байесовская оценка сходится к оценке максимального правдоподобия [см. (3.208)]. Конечно, формула (6.88) является частным случаем (6.89), когда шум белый и функции $s_j(t)$ ортогональны.

6.2.3. Оценка стационарного случайного сигнала на фоне шума. Рассмотрим узкополосный сигнал, представляющий нормальный стационарный случайный процесс с нулевым средним и корреляционной функцией $B_s(\tau)$. Запишем выражение этого сигнала через комплексную огибающую (см. § 5.2.2)

$$s(t) = \operatorname{Re} z_s(t)\, e^{i\omega_0 t} \qquad (6.90)$$

и воспользуемся ортогональным разложением этой огибающей [см. (3.36)]

$$z_s(t) = \sum_{k=1}^{\infty} \vartheta_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}, \qquad |t| \leqslant T, \qquad (6.91)$$

*) Аналогичными асимптотическими свойствами байесовские оценки обладают не только для квадратичной функции потерь, но и для функций потерь более общего вида [14].

где λ_k и $\varphi_k(t)$ — собственные числа и собственные функции линейного однородного интегрального уравнения

$$\varphi(t) = \lambda \int_{-T}^{T} B_{z_s}(t-y)\,\varphi(y)\,dy, \qquad (6.92)$$

и $B_{z_s}(\tau)$ — комплексная огибающая корреляционной функции сигнала. Координаты ϑ_k комплексной огибающей сигнала

$$\vartheta_k = \sqrt{\lambda_k} \int_{-T}^{T} z_s(t)\,\overline{\varphi_k(t)}\,dt, \qquad k = 1, 2, \ldots, \qquad (6.92')$$

некоррелированны. В силу нормального распределения действительная α_k и мнимая β_k части $\vartheta_k = \alpha_k + i\beta_k$ также и нормальны, причем [см. (5.38) — (5.38″)]

$$m_1\{\alpha_k\} = m_1\{\beta_k\} = m_1\{\alpha_k\beta_l\} = 0 \qquad (6.93)$$

для любых k и l,

$$m_1\{\alpha_k\alpha_l\} = m_1\{\beta_k\beta_l\} = 0, \qquad k \neq l, \qquad (6.93')$$

$$m_1\{\alpha_k^2\} = m_1\{\beta_k^2\} = \frac{1}{2}. \qquad (6.93'')$$

Задача состоит в определении байесовских оценок координат ϑ_k (или α_k и β_k), когда наблюдается на интервале $(-T, T)$ реализация $x(t)$ суммы сигнала $s(t)$ и стационарного узкополосного нормального шума с корреляционной функцией $B(\tau)$.

Ограничиваясь m первыми членами разложения (6.91) и вводя комплексную огибающую реализации $x(t) = \operatorname{Re} z(t)\, e^{i\omega_0 t}$, запишем функционал отношения правдоподобия для реализации $z(t)$ при фиксированных $\vartheta_1, \ldots, \vartheta_m$ [см. (5.54)]:

$$l[z(t) \mid \vartheta_1, \ldots, \vartheta_m] = \exp\left[\operatorname{Re} \int_{-T}^{T} V(t; \vartheta_1, \ldots, \vartheta_m) \times \right.$$

$$\left. \times \overline{z(t)}\,dt\right] \exp\left[-\frac{1}{2} \operatorname{Re} \sum_{k=1}^{m} \frac{\vartheta_k}{\sqrt{\lambda_k}} \int_{-T}^{T} \overline{V(t;\ \vartheta_1, \ldots, \vartheta_m)} \times \right.$$

$$\left. \times\, \varphi_k(t)\,dt\right], \qquad (6.94)$$

где $V(t;\ \vartheta_1, \ldots, \vartheta_m)$ определяется из неоднородного линейного интегрального уравнения

$$\int\limits_{-T}^{T} B_z(t-u)\, V(u;\ \vartheta_1, \ldots, \vartheta_m)\, du = \sum_{k=1}^{m} \vartheta_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}, \qquad |t| \leqslant T. \tag{6.94'}$$

Уравнение (6.94) можно заменить системой уравнений [см. (6.79)]

$$\int\limits_{-T}^{T} B_z(t-u)\, V_i(u)\, du = \frac{\varphi_i(t)}{\sqrt{\lambda_i}}, \qquad |t| \leqslant T, \tag{6.95}$$

причем

$$V(u;\ \vartheta_1, \ldots, \vartheta_m) = \sum_{j=1}^{m} \vartheta_j V_j(t). \tag{6.96}$$

Подставляя (6.96) в (6.94), перепишем функционал отношения правдоподобия в виде

$$l[z(t)\,|\,\vartheta_1, \ldots, \vartheta_m] = \exp\left[\mathrm{Re} \sum_{j=1}^{m} \vartheta_j \int\limits_{-T}^{T} V_j(t)\, \overline{z(t)}\, dt\right] \times$$
$$\times \exp\left[-\frac{1}{2} \mathrm{Re} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\vartheta_i \overline{\vartheta_j}}{\sqrt{\lambda_i}} \int\limits_{-T}^{T} \overline{V_j(t)}\, \varphi_i(t)\, dt\right]. \tag{6.97}$$

Введем следующие обозначения:

$$x_{Tj} = \mathrm{Re} \int\limits_{-T}^{T} V_j(t)\, \overline{z(t)}\, dt, \tag{6.98}$$

$$y_{Tj} = \mathrm{Im} \int\limits_{-T}^{T} V_j(t)\, \overline{z(t)}\, dt, \tag{6.98'}$$

$$s_{Tij} = \mathrm{Re} \int\limits_{-T}^{T} \overline{V_j(t)}\, \varphi_i(t)\, dt, \tag{6.99}$$

$$l_{Tij} = \mathrm{Im} \int\limits_{-T}^{T} \overline{V_j(t)}\, \varphi_i(t)\, dt. \tag{6.99'}$$

Используя указанные обозначения, из (6.97) получим

$$l\,[z(t)\,|\,\alpha_1,\ \beta_1,\ \ldots,\ \alpha_m,\ \beta_m] = \exp\left\{\sum_{j=1}^{m}(\alpha_j x_{Tj} - \beta_j y_{Tj}) - \right.$$

$$- \frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m}\frac{1}{\sqrt{\lambda_k}}\,[(\alpha_k\alpha_j + \beta_k\beta_j)\,s_{Tkj} -$$

$$\left. - (\beta_k\alpha_j - \alpha_k\beta_j)\,l_{Tkj}]\right\}, \qquad (6.100)$$

где $\alpha_k + i\beta_k = \vartheta_k,\ k = 1,\ \ldots,\ m.$

Из (6.93) — (6.93″) следует, что $2m$ координат сигнала $\alpha_1,\ \beta_1,\ \ldots,\ \alpha_m,\ \beta_m$ представляют совокупность $2m$ независимых нормальных случайных величин, для каждой из которых среднее равно нулю, а дисперсия равна 1/2. Таким образом,

$$w_{2m}(\alpha_1, \beta_1, \ldots, \alpha_m, \beta_m) = \frac{1}{(\pi)^m}\,e^{-\sum\limits_{k=1}^{m}(\alpha_k^2+\beta_k^2)}. \qquad (6.101)$$

Подставляя (6.100) и (6.101) в (6.82), получим байесовские оценки координат сигнала $s(t)$ [см. (6.90)] при квадратичной функции потерь

$$\hat{\alpha}_k = \frac{1}{K_m}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\alpha_k \exp\left\{-\sum_{j=1}^{m}(\alpha_j^2 + \beta_j^2) + \right.$$

$$+ \sum_{j=1}^{m}(\alpha_j x_{Tj} - \beta_j y_{Tj}) - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{1}{\sqrt{\lambda_i}}\,[(\alpha_i\alpha_j + \beta_i\beta_j)\,s_{Tij} -$$

$$\left. - (\beta_i\alpha_j - \alpha_i\beta_j)\,l_{Tij}]\right\} d\alpha_1\,d\beta_1\,\ldots\,d\alpha_m\,d\beta_m, \qquad (6.102)$$

$$\hat{\beta}_k = \frac{1}{K_m}\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\beta_k \exp\left\{-\sum_{j=1}^{m}(\alpha_j^2 + \beta_j^2) + \right.$$

$$+ \sum_{j=1}^{m}(\alpha_j x_{Tj} - \beta_j y_{Tj}) - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{1}{\sqrt{\lambda_i}}\,[(\alpha_i\alpha_j + \beta_i\beta_j)\,s_{Tij} -$$

$$\left. - (\beta_i\alpha_j - \alpha_i\beta_j)\,l_{Tij}]\right\} d\alpha_1\,d\beta_1\,\ldots,\,d\alpha_m\,d\beta_m, \qquad (6.102')$$

где

$$K_m = \int\limits_{-\infty}^{\infty} \ldots \int\limits_{-\infty}^{\infty} \exp \Big\{ -\sum_{j=1}^{m} (\alpha_j^2 + \beta_j^2) + \sum_{j=1}^{m} (\alpha_j x_{Tj} - \beta_j y_{Tj}) -$$

$$- \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{1}{\sqrt{\lambda_i}} [(\alpha_i \alpha_j + \beta_i \beta_j) s_{Tij} + (\beta_i \alpha_j - \alpha_i \beta_j) l_{Tij}] \Big\} \times$$

$$\times d\alpha_1 \, d\beta_1 \ldots d\alpha_m \, d\beta_m. \qquad (6.103)$$

Если аддитивный шум белый, то из (6.95) следует, что

$$V_i(t) = \frac{1}{N_0 \sqrt{\lambda_i}} \varphi_i(t), \quad i = 1, \ldots, m, \qquad (6.104)$$

где N_0 – спектральная плотность шума. Тогда из (6.98) и (6.98′) находим

$$x_{Tj} = \frac{1}{N_0 \sqrt{\lambda_j}} \operatorname{Re} \int\limits_{-T}^{T} \varphi_j(t) \overline{z(t)} \, dt, \qquad (6.105)$$

$$y_{Tj} = \frac{1}{N_0 \sqrt{\lambda_j}} \operatorname{Im} \int\limits_{-T}^{T} \varphi_j(t) \overline{z(t)} \, dt, \qquad (6.105')$$

а из (6.99), (6.99′) в силу ортогональности и нормировки собственных функций на интервале наблюдения $(-T, T)$

$$s_{Tii} = \frac{1}{N_0 \sqrt{\lambda_i}} \int\limits_{-T}^{T} |\varphi_i(t)|^2 \, dt = \frac{1}{N_0 \sqrt{\lambda_i}}, \qquad (6.106)$$

$$s_{Tij} = 0, \quad i \neq j, \quad l_{Tij} = 0 \text{ для любых } i \text{ и } j. \qquad (6.106')$$

В этом случае выражения (6.102)—(6.103) значительно упрощаются, причем переменные интегрирования разделяются. Вычислим сначала интеграл K_m:

$$K_m = \prod_{j=1}^{m} \int\limits_{-\infty}^{\infty} \exp \left(-\alpha_j^2 + \alpha_j x_{Tj} - \frac{\alpha_j^2}{2\lambda_j N_0} \right) d\alpha_j \times$$

$$\times \sum_{j=1}^{m} \int\limits_{-\infty}^{\infty} \exp \left(-\beta_j^2 + \beta_j y_{Tj} - \frac{\beta_j^2}{2\lambda_j N_0} \right) d\beta_j =$$

$$= \prod_{j=1}^{m} \frac{2\pi}{2 + \frac{1}{\lambda_j N_0}} \exp \left[- \frac{x_{Tj}^2 + y_{Tj}^2}{2 \left(2 + \frac{1}{\lambda_j N_0} \right)} \right]. \qquad (6.107)$$

Легко видеть, что в рассматриваемом частном случае белого шума в силу разделения переменных интегрирования

в (6.102) выражение оценки $\hat{\alpha}_k$ может быть представлено в виде

$$\hat{\alpha}_k = \frac{K'_{m-1}}{K_m} \exp\left[\frac{-x_{Tk}^2}{2\left(2+\frac{1}{\lambda_k N_0}\right)}\right] \int_{-\infty}^{\infty} \alpha_k \exp\left[-\frac{2+\frac{1}{\lambda_k N_0}}{2} \times \right.$$

$$\left. \times \left(\alpha_k - \frac{x_{Tk}}{2+\frac{1}{\lambda_k N_0}}\right)^2\right] d\alpha_k, \qquad (6.108)$$

где K'_{m-1} получается из (6.107), если в этой формуле опустить один сомножитель $\sqrt{2}\pi\left(2+\frac{1}{\lambda_k N_0}\right)^{-1/2} \exp\left[\frac{-x_{Tk}^2}{2(2+1/\lambda_k N_0)}\right]$. Наконец, вычисляя интеграл в (6.108), после очевидных сокращений получим

$$\hat{\alpha}_k = \frac{x_{Tk}}{2+\frac{1}{\lambda_k N_0}}, \qquad k=1, \ldots, m. \qquad (6.109)$$

Аналогично

$$\hat{\beta}_k = \frac{y_{Tk}}{2+\frac{1}{\lambda_k N_0}}, \qquad k=1, \ldots, m. \qquad (6.109')$$

Из (6.109) и (6.109′), учитывая (6.98), (6.98′) и (6.104), находим

$$\hat{\vartheta}_k = \frac{1}{2+\frac{1}{\lambda_k N_0}} \int_{-T}^{T} V_j(t) \overline{z(t)}\, dt =$$

$$= \frac{\sqrt{\lambda_k}}{1+2N_0\lambda_k} \int_{-T}^{T} \varphi_k(t) \overline{z(t)}\, dt \qquad (6.110)$$

и далее, подставляя (6.110) в (6.91), получаем оценку комплексной огибающей сигнала

$$\hat{z}_s(t) = \lim_{m\to\infty} \sum_{k=1}^{m} \frac{1}{1+2N_0\lambda_k} \varphi_k(t) \int_{-T}^{T} \varphi_k(u) \overline{z(u)}\, du. \qquad (6.111)$$

Вводя функцию

$$h(t, u) = \sum_{k=1}^{\infty} \frac{\varphi_k(t)\varphi_k(u)}{1+2N_0\lambda_k}, \qquad (6.112)$$

запишем (6.111) в виде

$$\hat{z}_s(t) = \int_{-T}^{T} h(t, u)\overline{z(u)}\,du, \qquad (6.113)$$

откуда находим байесовскую оценку сигнала

$$\hat{s}(t) = \operatorname{Re}\hat{z}_s(t)\,e^{i\omega_0 t} =$$

$$= \operatorname{Re} e^{i\omega_0 t}\int_{-T}^{T} h(t, u)\overline{z(u)}\,du, \quad |t| \leqslant T. \qquad (6.114)$$

В выражении (6.114) функция $h(t, u)$ определяется априорными сведениями о корреляционной функции сигнала и спектральной плотности шума, а $z(u)$ — комплексная огибающая реализации на интервале наблюдения.

6.2.4. Оценки случайных процессов, модулирующих высокочастотную несущую, на фоне аддитивного белого шума. Предположим теперь, что сигнал $s(t)$ представляет гармоническую несущую частоты ω_0, модулированную по амплитуде и по фазе (или по частоте) случайными процессами $a(t)$ и $\psi(t)$. Задача состоит в том, чтобы оценить значения этих модулирующих процессов по реализации суммы сигнала и аддитивного нормального шума, наблюдаемой на интервале $(-T, T)$. Чтобы не делать изложение слишком громоздким, ограничимся исследованием двух частных случаев, когда сигнал представляет амплитудно-модулированную или фазомодулированную несущую, хотя используемый метод решения задачи может быть использован и в более общих случаях.

Случай амплитудно-модулированного сигнала сводится к задаче, рассмотренной в § 6.2.2. Действительно, пусть

$$s(t) = a(t)\cos(\omega_0 t + \varphi_0). \qquad (6.115)$$

Воспользовавшись ортогональным разложением модулирующей функции

$$a(t) = \sum_{k=1}^{\infty} a_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}},$$

где λ_k и $\varphi_k(t)$ — собственные числа и собственные функции линейного интегрального уравнения

$$\varphi(t) = \lambda\int_{-T}^{T} B_a(t-y)\,\varphi(y)\,dy, \qquad |t| \leqslant T,$$

представим $s(t)$ в виде суммы, не отличающейся по виду от (6.77),

$$s(t) = \sum_{j=1}^{\infty} a_j s_j(t), \qquad (6.116)$$

причем

$$s_j(t) = \frac{\varphi_j(t)}{\sqrt{\lambda_j}} \cos(\omega_0 t + \varphi_0). \qquad (6.116')$$

Задача сводится теперь к совместной оценке координат a_k по методике, указанной в § 6.2.2.

Рассмотрим более подробно тот случай, когда несущая $A_0 \cos \omega_0 t$ модулирована по фазе стационарным нормальным случайным процессом $\psi(t)$ с нулевым средним и корреляционной функцией $B_\psi(\tau)$, медленно меняющейся по сравнению с $\cos \omega_0 \tau$, а аддитивный нормальный шум — белый со спектральной интенсивностью N_0 (и не зависит от сигнала).

Если воспользоваться ортогональным разложением модулирующей функции на интервале наблюдения

$$\psi(t) = \sum_{j=1}^{\infty} \psi_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}, \qquad |t| \leqslant T, \qquad (6.117)$$

где λ_k и $\varphi_k(t)$ — собственные числа и собственные функции линейного интегрального уравнения

$$\varphi(t) = \lambda \int_{-T}^{T} B_\psi(t-y)\varphi(y)\,dy, \qquad |t| \leqslant T, \quad (6.117')$$

то сигнал, представляющий модулированную по фазе несущую, можно представить в виде

$$s(t) = A_0 \cos\left[\omega_0 t + k_\text{ф} \sum_{k=1}^{\infty} \psi_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}\right], \qquad (6.118)$$

где ψ_k — независимые нормальные случайные величины с нулевыми средними и единичными дисперсиями. Константа $k_\text{ф}$ определяет наклон линейной характеристики фазового модулятора. Используя предположение, что аддитивный нормальный шум белый, запишем для рассматриваемого случая логарифм функционала отношения правдо-

подобия [см. (6.78) и (6.78′)], когда наблюдается реализация $x(t)$ на интервале $(-T, T)$:

$$\ln l[x(t) \mid \psi_1, \ldots, \psi_m, \ldots] = \frac{1}{N_0} \int_{-T}^{T} s(t) \times$$

$$\times \left[x(t) - \frac{1}{2} s(t)\right] dt = \frac{A_0}{N_0} \int_{-T}^{T} x(t) \times$$

$$\times \cos\left[\omega_0 t + k_ф \sum_{j=1}^{\infty} \psi_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}\right] dt - \frac{1}{2N_0} \int_{-T}^{T} s^2(t)\, dt,$$

и, пренебрегая в последнем слагаемом интегралом от быстро изменяющейся функции $\cos[2\omega_0 t + 2\psi(t)]$, получим

$$\ln l[x(t) \mid \psi_1, \ldots, \psi_m, \ldots] = \frac{A_0}{N_0} \int_{-T}^{T} x(t) \times$$

$$\times \cos\left[\omega_0 t + k_ф \sum_{j=1}^{\infty} \psi_k \frac{\varphi_k(t)}{\sqrt{\lambda_k}}\right] dt - \frac{A_0^2 T}{2N_0}. \qquad (6.119)$$

Тогда байесовские оценки координат ψ_k модулирующей функции при квадратичной функции потерь могут быть записаны следующим образом *):

$$\hat{\psi}_k = \left[\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \psi_k e^{-\frac{1}{2}\sum_{j=1}^{\infty}\psi_j^2} \exp\left\{\frac{A_0}{N_0} \int_{-T}^{T} x(t) \times\right.\right.$$

$$\left.\left.\times \cos\left[\omega_0 t + k_ф \sum_{j=1}^{\infty} \psi_j \frac{\varphi_j(t)}{\sqrt{\lambda_j}}\right] dt\right\} d\psi_1 \ldots d\psi_m \ldots\right] \times \qquad (6.120)$$

$$\times \left[\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\sum_{j=1}^{\infty}\psi_j^2} \exp\left\{\frac{A_0}{N_0} \int_{-T}^{T} x(t) \times\right.\right.$$

$$\left.\left.\times \cos\left[\omega_0 t + k_ф \sum_{j=1}^{\infty} \psi_j \frac{\varphi_j(t)}{\sqrt{\lambda_j}}\right] dt\right\} d\psi_1 \ldots d\psi_m \ldots\right]^{-1}.$$

---

*) В (6.120) и далее следовало бы брать конечное число N переменных интегрирования ψ_j и после интегрирования вычислять пределы при $N \to \infty$. Использованная сокращенная запись не влияет на окончательный результат.

Аналитическое исследование выражения (6.120) в общем виде не представляется возможным. Рассмотрим сначала случай, когда $\frac{2TA_0^2}{N_0} \ll 1$ (слабый сигнал). В этом случае можно пренебречь членами порядка $\left(\frac{2TA_0^2}{N_0}\right)^2$ и выше. Тогда, разлагая экспоненты в ряд по $\frac{A_0^2T}{N_0}$ и сохраняя только первые два члена, получим

$$\hat{\psi} = \Big[\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \psi_k e^{-\frac{1}{2}\sum_{j=1}^{\infty} \psi_j^2} \Big\{1 + \frac{A_0}{N_0} \int_{-T}^{T} x(t) \times$$

$$\times \cos\left[\omega_0 t + k_\phi \sum_{j=1}^{\infty} \psi_j \frac{\varphi_j(t)}{\sqrt{\lambda_j}}\right] dt \Big\} d\psi_1 \dots d\psi_m \dots \Big] \times$$

$$\times \Big[\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\sum_{j=1}^{\infty} \psi_j^2} \Big\{1 + \frac{A_0}{N_0} \int_{-T}^{T} x(t) \times$$

$$\times \cos\left[\omega_0 t + k_\phi \sum_{j=1}^{\infty} \psi_j \frac{\varphi_j(t)}{\sqrt{\lambda_j}}\right] dt \Big\} d\psi_1 \dots d\psi_m \Big]^{-1}$$

Так как при малом x

$$\frac{1}{1+x} = 1 - x + 0(x^2)$$

и

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \psi_k e^{-\frac{1}{2}\sum_{j=1}^{\infty} \psi_j^2} d\psi_1 \dots d\psi_m \dots = 0,$$

то из (6.120) с точностью до членов $0\left(\frac{A_0^4T^2}{N_0^2}\right)$

$$\hat{\psi}_k = \frac{A_0}{N_0} \operatorname{Re} \int_{-T}^{T} x(t)\, e^{i\omega_0 t} \left(\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \psi_k \prod_{j=1}^{\infty} \frac{1}{\sqrt{2\pi}} \times \right.$$

$$\left. \times \exp\left\{ -\frac{1}{2}\left[\psi_j^2 - 2k_\phi \frac{\varphi_j(t)}{\sqrt{\lambda_j}} \psi_j\right]\right\} d\psi_1 \dots d\psi_m \dots \right) dt,$$

и, вычисляя интеграл по переменным ψ_j, получим

$$\hat{\psi}_k = \frac{A_0 k_\Phi}{N_0 \sqrt{\lambda_k}} \int_{-T}^{T} \varphi_k(t)\, x(t)\, e^{k_\Phi^2 \sum_{j=1}^{\infty} \frac{\varphi_j^2(t)}{\lambda_j}} \cos\omega_0 t\, dt. \quad (6.121)$$

Далее, так как [см. (3.24′)]

$$\sum_{j=1}^{\infty} \frac{\varphi_j^2(t)}{\lambda_j} = B_\psi(0) = \sigma_\psi^2, \quad (6.122)$$

где σ_ψ^2 — дисперсия модулирующего процесса, то из (6.121) находим

$$\hat{\psi}_k = \frac{A_0 k_\Phi}{N_0 \sqrt{\lambda_k}} e^{k_\Phi^2 \sigma_\psi^2} \int_{-T}^{T} \varphi_k(t)\, x(t) \cos\omega_0 t\, dt. \quad (6.123)$$

Подставляя (6.123) в (6.117) и используя (3.24), получаем байесовскую оценку модулирующего процесса

$$\hat{\psi}(t) = \frac{A_0 k_\Phi}{N_0} e^{k_\Phi^2 \sigma_\psi^2} \int_{-T}^{T} B_\psi(t-u)\, x(u) \cos\omega_0 u\, du, \qquad |t| \leqslant T. \quad (6.124)$$

Из (6.124) следует, что

$$m_1\{\hat{\psi}(t)\} = 0. \quad (6.125)$$

Найдем корреляционную функцию оценки $\hat{\psi}(t)$. Так как в силу независимости сигнала и шума

$$B_x(u, v) = B_s(u, v) + N_0 \delta(v-u),$$

то

$$B_{\hat{\psi}}(u, v) = m_1\{\hat{\psi}(u)\hat{\psi}(v)\} =$$

$$= \frac{A_0^2 k_\Phi^2}{N_0^2} e^{2k_\Phi^2 \sigma_\psi^2} \int_{-T}^{T}\int_{-T}^{T} B_\psi(u-x)\, B_\psi(v-y) \times$$

$$\times [B_s(x, y) + N_0 \delta(x-y)] \cos\omega_0 x \cos\omega_0 y\, dx\, dy.$$

Если пренебречь интегралами от слагаемых, содержащих $\cos\omega t$ при $\omega \geqslant \omega_0$ (как это уже было сделано выше), то, учитывая, что корреляционная функция $B_s(x, y)$ модулированной несущей пропорциональна указанным быстроменяющимся функциям (см. § 12.1.2 в первой книге),

получим

$$B_{\hat{\psi}}(u, v) = \frac{A_0^2 k_\Phi^2}{N_0^2} e^{2k_\Phi^2 \sigma_\psi^2} \int_{-T}^{T} B_\psi(u-y) B_\psi(v-y)\, dy. \quad (6.126)$$

Дисперсия оценки

$$\sigma_{\hat{\psi}}^2(t) = B_{\hat{\psi}}(t, t) = \frac{A_0^2 k_\Phi^2}{N_0^2} e^{2k_\Phi^2 \sigma_\psi^2} \int_{-T}^{T} B_\psi^2(u-y)\, dy \quad (6.127)$$

или

$$\left[\frac{\sigma_{\hat{\psi}}(t)}{\sigma_\psi}\right]^2 = \frac{A_0^2 k_\Phi^2 \sigma_\psi^2}{2N_0} e^{2k_\Phi^2 \sigma_\psi^2} \int_{-T}^{T} R_\psi^2(\tau)\, d\tau, \quad (6.127')$$

где $R_\psi(\tau)$ — коэффициент корреляции модулирующего процесса.

При $T \to \infty$

$$\left[\frac{\sigma_{\hat{\psi}}(t)}{\sigma_\psi}\right]^2 \sim \frac{A_0^2 k_\Phi^2 \sigma_\psi^2}{N_0} e^{2k_\Phi^2 \sigma_\psi^2} \int_{0}^{\infty} R_\psi^2(\tau)\, d\tau. \quad (6.128)$$

Задачи

6.1. Показать, что если распределение амплитуды узкополосного сигнала — релеевское

$$w_1(a) = \frac{2a}{\sigma_a^2} e^{-\frac{a^2}{\sigma_a^2}}, \quad a > 0,$$

и распределение независимой от амплитуды фазы равномерное, то при квадратичной функции потерь байесовская оценка амплитуды в аддитивном нормальном шуме равна

$$\hat{a} = \sqrt{\frac{\pi}{2d_T^2 + \frac{4}{\sigma_a^2}}}\; {}_1F_1\left[-\frac{1}{2};\; 1;\; -\frac{r_T^2}{2\left(d_T^2 + \frac{2}{\sigma_a^2}\right)}\right], \quad (1)$$

где ${}_1F_1$ — гипергеометрическая функция (см. приложение V в первой книге), а величины r_T^2 и d_T^2 определены согласно (6.70) и (6.71).

6.2. Решить задачу § 6.2.2 в предположении, что результаты наблюдений представлены не реализацией $x(t) = s(t;\ \vartheta_1, \ldots, \vartheta_m) + \xi(t)$, а дискретной выборкой $x = (x_1, \ldots, x_N)$, где $x_i = x(t_i)$, t_i — моменты времени, принадлежащие интервалу наблюдения $(-T, T)$. Показать, что байесовские оценки параметров ϑ_k могут быть пред-

ставлены в векторной форме

$$\hat{\boldsymbol{\vartheta}} = \frac{\int\limits_{G_\vartheta} \boldsymbol{\vartheta} w(\boldsymbol{\vartheta}) \exp\left\{-\frac{N}{2\,\mathrm{tr}\,\mathbf{K}} Q_N(\mathbf{x}-\mathbf{s}\boldsymbol{\vartheta},\ \mathbf{k})\right\} d\boldsymbol{\vartheta}}{\int\limits_{G_\vartheta} w(\boldsymbol{\vartheta}) \exp\left\{-\frac{N}{2\,\mathrm{tr}\,\mathbf{K}} Q_N(\mathbf{x}-\mathbf{s}\boldsymbol{\vartheta},\ \mathbf{k})\right\} d\boldsymbol{\vartheta}}, \tag{2}$$

где G_ϑ — область пространства параметров, в которой задано распределение $w(\boldsymbol{\vartheta})$; $\mathbf{s}$ — матрица размером $N \times m$ с линейно-независимыми (в алгебраическом смысле) столбцами, причем j-й столбец матрицы представляет вектор с компонентами $s_j(t_1), \ldots, s_j(t_N)$, $j=1, \ldots, m$; $\mathbf{k}$ — нормированная корреляционная матрица случайного слагаемого $\xi(t)$;

$$\mathbf{k} = \frac{N\mathbf{K}}{\mathrm{tr}\,\mathbf{K}}, \tag{2'}$$

$\mathrm{tr}\,\mathbf{K}$ — след корреляционной матрицы $\mathbf{K}$,

$$Q_N(\mathbf{x}-\mathbf{s}\boldsymbol{\vartheta},\ \mathbf{k}) = (\mathbf{x}-\mathbf{s}\boldsymbol{\vartheta})'\,\mathbf{k}^{-1}(\mathbf{x}-\mathbf{s}\boldsymbol{\vartheta}). \tag{3}$$

6.3. Получить следующие формулы для элементов информационной матрицы оценок амплитуды и фазы узкополосного сигнала в аддитивном белом шуме:

$$I_T^{(11)}(A,\ \varphi) = m_1\left\{\left(\frac{\partial}{\partial A}\ln l\,[x(t)\mid A,\ \varphi]\right)^2\right\} =$$

$$= m_1\left\{\frac{16T^2}{N_0^2}\left[\left(\frac{1}{2T}\int\limits_{-T}^{T} A(t)\,dt\right)\cos\varphi + \left(\frac{1}{2T}\int\limits_{-T}^{T} C(t)\,dt\right)\sin\varphi -\right.\right.$$

$$\left.\left. - A\right]^2\right\} = \frac{8T}{N_0} = \frac{4d_T^2}{A^2}, \tag{4}$$

$$I_T^{(22)}(A,\ \varphi) = m_1\left\{\left(\frac{\partial}{\partial\varphi}\ln l\,[x(t)\mid A,\ \varphi]\right)^2\right\} =$$

$$= m_1\left\{\frac{16T^2}{N_0^2}\left[\left(\frac{1}{2T}\int\limits_{-T}^{T} A(t)\,dt\right)\sin\varphi -\right.\right.$$

$$\left.\left. - \left(\frac{1}{2T}\int\limits_{-T}^{T} (t)\,dt\right)\cos\varphi\right]^2\right\} = \frac{8TA^2}{N_0} = 4d_T^2, \tag{5}$$

$$I_T^{(12)}(A,\ \varphi) = m_1\left\{\frac{\partial}{\partial A}\ln l\,[x(t)\mid A,\ \varphi]\times\right.$$

$$\left.\times\frac{\partial}{\partial\varphi}\ln l\,[x(t)\mid A,\ \varphi]\right\} = 0, \tag{6}$$

где

$$d_T^2 = \frac{2TA^2}{N_0}. \tag{6'}$$

Из (4) – (6) найти элементы матрицы, обратной информационной:

$$Y_T^{(11)}(A,\ \varphi)=\frac{1}{4d_T^2}, \tag{7}$$

$$Y_T^{(22)}(A,\ \varphi)=\frac{A^2}{4d_T^2}, \tag{8}$$

$$Y_T^{(12)}(A,\ \varphi)=0, \tag{9}$$

и доказать асимптотическую эффективность при $d_T\to\infty$ совместных оценок максимального правдоподобия амплитуды и фазы.

6.4. Доказать, что для узкополосного сигнала с гауссовой огибающей $u(t)=e^{-\alpha^2t^2}$ и линейной частотной модуляцией $\omega(t)=\omega_0+\lambda t$, т. е. для сигнала с нормированной комплексной огибающей вида

$$a(t)=\frac{4\alpha T}{\sqrt{2\pi}}e^{-(\alpha^2-i\lambda)t^2},\quad \alpha T\gg 1, \tag{10}$$

модуль функции неопределенности равен

$$|\Psi(\tau)|=e^{-\frac{\alpha^2}{2}m^2\tau^2}, \tag{11}$$

где m — так называемый коэффициент сжатия, определяемый по формуле

$$m=\sqrt{1+\frac{\lambda^2}{\alpha^4}} \tag{12}$$

(при отсутствии частотной модуляции $m=1$).

Используя (11), показать, что дисперсия оценки времени прихода сигнала, смешанного с аддитивным нормальным белым шумом, когда отношение d_T^2 энергии сигнала к спектральной плотности шума велико, равна

$$M_2\{\hat{\tau}\}=\frac{1}{d_T^2\alpha^2m^2}. \tag{13}$$

6.5. Доказать, что при совместном измерении времени прихода и допплеровского смещения частот модуль двумерной функции неопределенности для сигнала вида (10) задачи 6.4 равен

$$|\Psi(\tau,\ \Omega)|=e^{-\frac{1}{2}\left(\alpha^2m^2\tau^2-\frac{\lambda}{\alpha^2}\Omega\tau+\frac{\Omega^2}{4\alpha^2}\right)}. \tag{14}$$

Используя (14), показать, что дисперсии оценок времени прихода и допплеровского смещения частот и коэффициент корреляции этих оценок, когда отношение энергии сигнала к спектральной плотности белого шума $d_T^2\gg 1$, определяются по формулам

$$M_2\{\hat{\tau}\}=\frac{1}{d_T^2\alpha^2}, \tag{15}$$

$$M_2\{\Omega\} = \frac{4\alpha^2 m^2}{d_T^2}, \tag{16}$$

$$r = \sqrt{1 - \frac{1}{m^2}}. \tag{17}$$

6.6. Предположим, что амплитуда и фаза узкополосного квазидетерминированного сигнала [вида (6.66)] независимы, причем априорное распределение амплитуды релеевское, а фазы—равномерное. Доказать, что байесовская оценка амплитуды равна

$$\hat{a} = \frac{\sigma_a}{\sqrt{1+d_T^2}} \sqrt{\frac{\pi}{2}} \left[\left(1 + \frac{\rho_T^2}{2}\right) I_0\left(\frac{\rho_T^2}{4}\right) + \frac{\rho_T^2}{2} I_1\left(\frac{\rho_T^2}{4}\right) \right] e^{-\frac{\rho_T^2}{4}}, \tag{18}$$

где

$$\rho_T = \frac{r_T}{\sqrt{1+d_T^2}}, \tag{19}$$

а величины r_T^2 и d_T^2 определяются по формулам (6.70) и (6.71).

Показать, что при $d_T \to \infty$.

$$\hat{a} \sim \frac{\sigma_a}{d_T^2} r_T. \tag{20}$$

Убедиться, что для белого шума оценка (20) совпадает с оценкой, определяемой по формуле (6.76).

ЛИТЕРАТУРА

МОНОГРАФИИ

1. Вакман Д. Е. Сложные сигналы и принцип неопределенности в радиолокации. Изд-во «Советское радио», 1965, гл. 2, 3.
2. Вудворд Ф. М. Теория вероятностей и теория информации с применениями в радиолокации. Пер. с англ., под ред. Г. С. Горелика. Изд-во «Советское радио», 1955, гл. 5, 6, 7.
3. Вопросы статистической теории радиолокации. Под ред. Г. П. Тартаковского, т. II. Изд-во «Советское радио», 1964.
4. Гуткин Л. С. Теория оптимальных методов радиоприема при флуктуационных помехах. Госэнергоиздат, 1961, гл. 13, 16.
5. Миддлтон Д. Введение в статистическую теорию связи, т. II. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1962, гл. 21.
6. Миддлтон Д. Очерки теории связи. Пер. с англ., под ред. Б. Р. Левина. Изд-во «Советское радио», 1966, гл. 3.
7. Фалькович С. Е. Прием радиолокационных сигналов на фоне флюктуационных помех. Изд-во «Советское радио», 1961, гл. 4, 5.

8. Харкевич А. А. Борьба с помехами. Изд-во «Наука», 1965, § 15, 16.
9. Хелстром К. Статистическая теория обнаружения сигналов. Пер. с англ., под ред. Ю. В. Кобзарева. Изд-во иностранной литературы, 1963, гл. 8.
10. Ширман Я. Д., Голиков В. Н. Основы теории обнаружения радиолокационных сигналов и измерения их параметров. Изд-во «Советское радио», 1963, гл. 6, 7.
11. Большаков И. А. Статистические проблемы многих сигналов в радиотехнике и радиолокации. Изд-во «Советское радио», 1968.

СТАТЬИ

12. Келли Е. Радиолокационное измерение дальности, скорости и ускорения. «Зарубежная радиоэлектроника», 1962, № 2, стр. 35—47.
13. Левин Б. Р., Шинаков Ю. С. Предельная форма байесовской оценки коэффициентов регрессии при нормальном нестационарном шуме. «Проблемы передачи информации», 1967, № 1.
14. Левин Б. Р., Шинаков Б. С. Некоторые асимптотические свойства байесовской оценки коэффициента регрессии в присутствии нестационарного шума. «Проблемы передачи информации», 1967, № 3.
15. Abbate J. V., Schilling D. L. Estimation of random Phase — and Frequency Modulating Signals using a Bayes Estimator. IEEE Trans., 1965, IT-11, № 3.
16. Kelly E. J., Reed I. S., Root W. L. The detection of radar echoes in noise, II. J. Soc. Ind. Appl. Math., 1960, v. 8, № 3.
17. Sibert W. M. A radar detection philosophy. IRE Trans., 1965, Sept. IT-2.

ГЛАВА СЕДЬМАЯ

ЭЛЕМЕНТЫ ТЕОРИИ КЛАССИФИКАЦИИ С ОБУЧЕНИЕМ

7.1. ХАРАКТЕРИСТИКА ПРОБЛЕМЫ КЛАССИФИКАЦИИ С ОБУЧЕНИЕМ

Проблема классификации результатов наблюдений была сформулирована в самом начале книги как проблема выбора из набора решений $\gamma_0, \ldots, \gamma_m$ решения γ_k о принадлежности совокупности результатов наблюдений $x_1, \ldots, x_n$ одному из $m+1$ заданных распределений $W_n(x_1, \ldots, x_n \mid s_k)$, $k = 0, \ldots, m$.

Использование тех или иных критериев качества для получения оптимального алгоритма классификации обуславливалось знанием априорных вероятностей принадлежности каждому из распределений и заданием функции потерь. При этом апостериорные (условные) функции распределения выборочных значений — функции правдоподобия $W_n(x_1, \ldots, x_n \mid s_k)$ предполагались *полностью известными*. В том случае, когда параметры этих функций оказывались случайными, предполагалось, что точно известны функции распределения этих параметров.

Рассмотренные в пятой главе задачи обнаружения и различения сигналов на фоне шума служат примерами задач синтеза оптимальных алгоритмов классификации по выборочным значениям наблюдаемого процесса или по непрерывной реализации этого процесса, для решения которых использовались байесовские критерии качества.

Далеко не всегда байесовский подход к проблеме синтеза оптимальных алгоритмов классификации может быть практически реализован из-за возникающих чисто математических трудностей, связанных с нахождением экстремумов достаточно сложных функционалов. Но даже, если оставить в стороне ограниченные возможности аппарата математического программирования, указанный подход встречается иногда с принципиальными трудностями. Это относится, прежде всего, к вероятностному описанию

изучаемого явления, которым исследователь может не располагать вовсе или которое может быть неполным. Так часто неизвестны параметры функции правдоподобия выборочных значений или априорное распределение этих параметров, а в других случаях — и сама функция правдоподобия. Произвол в выборе вида функции потерь является другой причиной критического отношения к байесовскому подходу.

Одним из путей, выводящих из тупика априорной недостаточности, является поиск алгоритмов, устойчивых к изменению вероятностных характеристик изучаемых случайных процессов. Некоторые возможности, частично уже реализованные, открываются при исследовании асимптотического поведения оптимальных байесовских алгоритмов, когда какой-либо параметр, например соответствующим образом определенное отношение сигнал/шум, мал или неограниченно возрастает. Эти предельные алгоритмы при определенных условиях оказываются инвариантными по отношению к априорным данным.

Более широкий горизонт открывается с вершин непараметрических методов. Некоторые из этих методов были упомянуты в первых главах книги. Таковы методы проверки гипотезы, базирующиеся на использовании упорядоченности выборочных значений (см. § 1.5.3) или критериев согласия (§ 2.6), при использовании которых сохраняется постоянным уровень значимости, каково бы ни было распределение, соответствующее проверяемой гипотезе. В частности, в задачах обнаружения сигналов в помехах непараметрические методы позволяют найти алгоритмы, обеспечивающие постоянство вероятности ложных тревог при изменении распределения вероятностей помех.

Другой путь преодоления трудностей, вызванных отсутствием априорной информации, указывает теория игр. Задачи классификации рассматриваются при этом как «игра» с противником (природой), стратегия которого, вообще говоря, неизвестна. Стратегия исследователя часто строится на предположении, что природа выбирает всегда наименее благоприятное для него распределение вероятностей исследуемого случайного процесса. Такая стратегия приводит к минимаксному правилу выбора решения (см. § 1.1.5). Байесовскую теорию решений можно рассматривать как специальный раздел теории игр для тех случаев, когда «игра» ведется против «слабого» противника,

стратегия которого в вероятностном смысле заранее известна исследователю.

Методы непараметрической статистики и теории игр здесь не рассматриваются. Настоящая глава посвящена третьему возможному пути решения задач классификации для тех ситуаций, когда функции распределения выборочных значений частично или полностью неизвестны. В этом случае классификации предшествует период *обучения*, т. е. формирования на основании наблюдений оценок неизвестных функций распределения или оценок параметров, если вид функций распределения задан. Эти оценки используются затем для классификации последующих результатов наблюдений вместо неизвестных истинных вероятностных характеристик исследуемых процессов.

Различают обучение с учителем, в результате которого получают последовательность эталонных наблюдений, причем относительно каждого из ее элементов известно какому из распределений он принадлежит (классифицированная обучающая выборка), и обучение без учителя (самообучение), при котором указанные выше оценки формируются по неклассифицированным обучающим выборкам. Иногда заранее ограничивают класс допустимых алгоритмов классификации, параметры которых оцениваются по выбранному критерию качества в процессе обучения или самообучения.

В приемном устройстве, предназначенном для классификации (например, для обнаружения или различения сигналов на фоне шума), когда условные функции распределения результатов наблюдения не определены, в период обучения или самообучения происходит изменение параметров или структуры алгоритма классификации с целью приближения их к оптимальным с точки зрения заданного критерия качества. Подобные приемные устройства принято называть адаптивными (или самоорганизующимися).

Часто при решении задач классификации используются не сами наблюдения, а некоторые функции (или функционалы) от этих наблюдений, называемые информативными признаками класса (или признаками образа). В указанных терминах рассматриваемая проблема относится к проблеме распознавания образов по выбранным информативным признакам. Выбор информативных признаков зависит от характера решаемой задачи. Общие подходы в этом на-

правлении еще только намечаются. Естественным является стремление решить задачу, имея по возможности меньшее число информативных признаков. В некоторых случаях в качестве информативных признаков могут быть выбраны достаточные статистики.

Таким образом, в наиболее общем виде система, реализующая алгоритм классификации, включает два устройства: приемное, преобразующее результаты наблюдений в информативные признаки, характеризующие образ, и решающее, которое относит сформированную на выходе совокупность признаков к одному из классов (образов).

Выборочные значения $x_1, \ldots, x_p$ могут рассматриваться или как результаты непосредственных наблюдений, или как информативные признаки.

7.2. КЛАССИФИКАЦИЯ В СЛУЧАЕ НОРМАЛЬНЫХ РАСПРЕДЕЛЕНИЙ

7.2.1. Неизвестные средние. Рассмотрим сначала простейший случай, когда в результате обучения с учителем получена классифицированная обучающая выборка, причем $x_1^{(0)}, \ldots, x_{n_0}^{(0)}$ принадлежит классу s_0, характеризуемому одномерным нормальным распределением $w_1(x \mid s_0)$, а $x_1^{(1)}, \ldots, x_{n_1}^{(1)}$ — классу s_1, характеризуемому одномерным нормальным распределением $w_1(x \mid s_1)$. Предположим сначала, что неизвестны только средние значения нормальных распределений, а дисперсии известны и имеют одинаковую величину σ^2. Примем в качестве оценок неизвестных средних оценки максимального правдоподобия по классифицированной обучающей выборке, т. е.

$$\hat{a}_0 = \frac{1}{n_0} \sum_{i=1}^{n_0} x_i^{(0)}, \quad \hat{a}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i^{(1)}. \tag{7.1}$$

Тогда, используя эти оценки вместо неизвестных средних, можно согласно § 1.3 сформулировать следующее оптимальное по критерию максимального правдоподобия правило классификации наблюдений $x_1, \ldots, x_n$: наблюдения относятся к классу s_1, если

$$\left(\frac{1}{n} \sum_{i=1}^{n} x_i - \frac{\hat{a}_0 + \hat{a}_1}{2}\right)(\hat{a}_1 - \hat{a}_0) > 0 \tag{7.2}$$

и к классу s_0, если выполняется неравенство, противоположное (7.2).

Формулу (7.2) можно переписать в виде системы неравенств

$$\frac{1}{n}\sum_{i=1}^{n} x_i \geqslant \frac{\hat{a}_0+\hat{a}_1}{2}, \quad \hat{a}_1 > \hat{a}_0 \tag{7.3}$$

или

$$\frac{1}{n}\sum_{i=1}^{n} x_i \leqslant \frac{\hat{a}_0+\hat{a}_1}{2}, \quad \hat{a}_1 < \hat{a}_0. \tag{7.3'}$$

Таким образом, алгоритм классификации сводится к вычислению среднего арифметического наблюдаемых значений и сравнению его с порогом, *зависящим от обучающих выборок*.

Левая часть выражения (7.2) представляет произведение коррелированных нормальных случайных величин

$$y = \hat{a}_1 - \hat{a}_0, \tag{7.4}$$

$$z = \frac{1}{n}\sum_{i=1}^{n} x_i - \frac{\hat{a}_0+\hat{a}_1}{2}. \tag{7.4'}$$

Средние и дисперсии этих величин равны

$$m_1\{y\} = a_1 - a_0, \quad M_2\{y \mid s_0, s_1\} = \sigma^2\left(\frac{1}{n_0}+\frac{1}{n_1}\right), \tag{7.5}$$

$$m_1\{z \mid s_1\} = \frac{a_1-a_0}{2} = -m_1\{z \mid s_0\}, \tag{7.6}$$

$$M_2\{z \mid s_1, s_0\} = \sigma^2\left(\frac{1}{n}+\frac{1}{4n_0}+\frac{1}{4n_1}\right), \tag{7.6'}$$

а коэффициент корреляции

$$R_{yz} = \frac{m_1\{yz\} - m_1\{y\}\, m_1\{z\}}{\sqrt{M_2\{y\}\, M_2\{z\}}} = \frac{\frac{1}{n_1}-\frac{1}{n_0}}{\sqrt{\left(\frac{1}{n}+\frac{1}{4n_0}+\frac{1}{4n_1}\right)\left(\frac{1}{n_0}+\frac{1}{n_1}\right)}}, \tag{7.7}$$

причем a_1 и a_0 — *неизвестные* средние нормальных распределений $w_1(x \mid a_1)$ и $w_1(x \mid a_0)$. Заметим, что при $n_0 = n_1$ случайные величины y и z не коррелированы.

Когда размеры обучающих выборок n_0 и n_1 неограниченно возрастают, оценки $\hat{a}_0$ и $\hat{a}_1$ сходятся по вероятности к средним значениям a_0 и a_1 соответственно (так как выборочное среднее является состоятельной оценкой среднего). Учитывая, кроме того, что пределы по вероятности сумм, разностей и произведений случайных величин равны суммам, разностям и произведениям пределов, находим из (7.3), что при неограниченном увеличении размеров обучающих выборок (и фиксированном размере n выборки, которую необходимо отнести к одному из двух указанных выше нормальных распределений) оптимальное правило классификации асимптотически сходится по вероятности к байесовскому правилу проверки простой гипотезы о среднем значении нормальной случайной величины [см. (1.67)]. Вероятности ошибок классификации стремятся при этом асимптотически к величинам α и β, определяемым по формулам (1.71) и (1.72).

Алгоритм классификации (7.3) обобщается на многомерный случай, когда решается задача о принадлежности наблюдаемой выборки $\mathbf{X}$ одному из двух p-мерных нормальных распределений с неизвестными векторами средних и заданными корреляционными матрицами $\mathbf{M}_1 = \mathbf{M}_2 = \mathbf{M}$. Пусть в результате обучения с учителем получена классифицированная обучающая выборка: из первого распределения $\mathbf{x}_1^{(0)}, \ldots, \mathbf{x}_{n_0}^{(0)}$ и из второго $\mathbf{x}_1^{(1)}, \ldots, \mathbf{x}_{n_1}^{(1)}$. Каждый элемент указанных выборок представляет p-мерный вектор. Аналогично (7.1) в качестве оценок неизвестных векторов средних принимаются средние арифметические

$$\hat{\mathbf{a}}_0 = \frac{1}{n_0} \sum_{i=1}^{n_0} \mathbf{x}_i^{(0)}, \quad \hat{\mathbf{a}}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{x}_i^{(1)}. \tag{7.8}$$

Алгоритм классификации, обобщающий (7.2), сводится к сравнению с порогом величины V, равной

$$V = \left(\mathbf{X} - \frac{\hat{\mathbf{a}}_0 + \hat{\mathbf{a}}_1}{2}\right)' \mathbf{M}^{-1} (\hat{\mathbf{a}}_1 - \hat{\mathbf{a}}_0). \tag{7.9}$$

Распределение статистики V изучалось в работах [24, 45]. Было показано [32], в частности, что статистика

$$Z = \frac{4 n_0 n_1 V}{\sqrt{(n_0 + n_1)(n_0 + n_1 + 4 n_0 n_1)}}$$

может быть приведена к линейной комбинации двух независимых случайных величин $\chi^2(p, \lambda_1)$ и $\chi^2(p, \lambda_2)$

$$Z = (1+\rho)\,\chi^2(p, \lambda_1) - (1-\rho)\,\chi^2(p, \lambda_2), \qquad (7.10)$$

где $\chi^2(p, \lambda)$ имеет нецентральное χ^2-распределение с p степенями свободы и параметром нецентральности λ, а величина ρ определяется по формуле

$$\rho = \frac{n_1 - n_0}{\sqrt{(n_0+n_1)(n_0+n_1+4n_0n_1)}}. \qquad (7.10')$$

Параметры нецентральности определяются из соотношений

$$\lambda_1 = \frac{n_0n_1}{4(1+\rho)}\left[\frac{1}{\sqrt{n_0+n_1}} \mp \frac{1}{\sqrt{n_0+n_1+4n_0n_1}}\right]^2 d, \qquad (7.11)$$

$$\lambda_2 = \frac{n_0n_1}{4(1-\rho)}\left[\frac{1}{\sqrt{n_0+n_1}} \pm \frac{1}{\sqrt{n_0+n_1+4n_0n_1}}\right]^2 d, \qquad (7.11')$$

где

$$d^2 = (\mathbf{a}_1 - \mathbf{a}_0)'\,\mathbf{M}^{-1}(\mathbf{a}_1 - \mathbf{a}_0). \qquad (7.12)$$

Верхний знак в квадратных скобках соответствует случаю, когда $m_1\{\mathbf{X}\} = \mathbf{a}_0$, а нижний — когда $m_1\{\mathbf{X}\} = \mathbf{a}_1$.

При $n_0 \to \infty$ и $n_1 \to \infty$ распределение V приближается к нормальному с параметрами $\left(\frac{1}{2}d^2, d^2\right)$, если $m_1\{\mathbf{X}\} = \mathbf{a}_0$, и с параметрами $\left(-\frac{1}{2}d^2, d^2\right)$, если $m_1\{\mathbf{X}\} = \mathbf{a}_1$.

7.2.2. Неизвестные средние (обучение без учителя). Вернемся к постановке задачи, изложенной в начале § 7.2.1, с той лишь разницей, что обучающая выборка $x_1, \ldots, x_n$ не классифицирована. Предполагая, что появление любого из двух классов s_1 и s_0 в каждом наблюдении априори равновероятно, можно каждый элемент обучающей выборки рассматривать как принадлежащий общему бимодальному распределению [27] (см. задачу 1.5)

$$w_1(x \mid a_1, a_0) = \frac{1}{2\sqrt{2\pi}\,\sigma}\left[e^{-\frac{(x-a_1)^2}{2\sigma^2}} + e^{-\frac{(x-a_0)^2}{2\sigma^2}}\right]. \qquad (7.13)$$

Среднее значение случайной величины, подчиняющейся распределению (7.13), равно

$$a = \frac{a_1 + a_0}{2}. \qquad (7.14)$$

Так как a_1 и a_0 неизвестны, то неизвестна и величина a. В качестве оценки среднего значения a распределения (7.13)

можно принять выборочное среднее, полученное из всей обучающей выборки

$$\hat{a} = \frac{1}{n} \sum_{i=1}^{n} x_i. \qquad (7.15)$$

Используя оценку (7.15) вместо неизвестного среднего a и принимая критерий минимума вероятности ошибочных классификаций (максимального правдоподобия), можно сформулировать следующее оптимальное правило классификации. Наблюдение X относится к классу s_1, если

$$X \geqslant \hat{a}, \qquad (7.16)$$

и к классу s_0, если выполняется неравенство, противоположное (7.16).

При неограниченном увеличении n сформулированное правило классификации асимптотически стремится к байесовскому правилу проверки простой гипотезы о среднем значении нормальной случайной величины.

Алгоритм классификации (7.16) обобщается на многомерный случай при сохранении сферической симметрии плотностей вероятности, когда решается задача о принадлежности наблюдаемой выборки $\mathbf{X}$ одному из двух p-мерных нормальных распределений с неизвестными векторами средних $\mathbf{a}_1$ и $\mathbf{a}_0$ и заданными корреляционными матрицами $\mathbf{M}_1 = \mathbf{M}_2 = \sigma^2\mathbf{I}$, где $\mathbf{I}$ — единичная матрица. В этом случае общее многомерное распределение двух классов имеет следующий вид:

$$w_p(\mathbf{X} \mid \mathbf{a}_1, \mathbf{a}_0) = \frac{1}{2(2\pi)^{\frac{p}{2}} \sigma^p} \left\{ \exp\left[-\frac{1}{2\sigma^2} (\mathbf{X} - \mathbf{a}_1)' (\mathbf{X} - \mathbf{a}_1) \right] + \right.$$

$$\left. + \exp\left[-\frac{1}{2\sigma^2} (\mathbf{X} - \mathbf{a}_0)' (\mathbf{X} - \mathbf{a}_0) \right] \right\} =$$

$$= \frac{1}{(2\pi)^{\frac{p}{2}} \sigma^p} \exp\left[-\frac{1}{2\sigma^2} \mathbf{b}'\mathbf{b} \right] \exp\left[-\frac{1}{2\sigma^2} (\mathbf{X} - \mathbf{a})' (\mathbf{X} - \mathbf{a}) \right] \times$$

$$\times \operatorname{ch} \frac{\mathbf{b}' (\mathbf{X} - \mathbf{a})}{\sigma^2}, \qquad (7.17)$$

где

$$\mathbf{a} = \frac{\mathbf{a}_1 + \mathbf{a}_0}{2}; \quad \mathbf{b} = \frac{\mathbf{a}_1 - \mathbf{a}_0}{2}. \qquad (7.17')$$

Вектор **a** является вектором средних значений распределения (7.17), а элементы корреляционной матрицы **M** этого распределения равны

$$M_{ij} = \int_{-\infty}^{\infty} \dots \int (x_i - a_i)(x_j - a_j) w_p(\mathbf{X} \mid \mathbf{a}, \mathbf{b})\, dx_1 \dots dx_p =$$

$$= b_i b_j + \sigma^2 \delta_{ij}, \qquad (7.18)$$

где $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$,

a_i, $i = 1, \dots, p$ — компоненты вектора **a**,

b_j, $j = 1, \dots, p$ — компоненты вектора **b**.

Если векторы средних $\mathbf{a}_1$ и $\mathbf{a}_0$ известны, то оптимальное байесовское правило разбиения выборочного пространства производит гиперплоскость, которая перпендикулярна линии, соединяющей точки $\mathbf{X} = \mathbf{a}_1$ и $\mathbf{X} = \mathbf{a}_0$, и делит эту линию пополам (см. задачу 1.6). Наблюдение **X** относится к одному или другому классу в зависимости от знака величины $\mathbf{b}'(\mathbf{X} - \mathbf{a})$ [ср. с (7.9)].

В том случае, когда векторы средних для обоих классов неизвестны, **a** и **b** следует заменить оценками, полученными при обучении. При самообучении по неклассифицированной выборке $\mathbf{x}_1, \dots, \mathbf{x}_n$ эти оценки получаются из выборочного среднего и выборочной корреляционной матрицы. Оценка вектора средних распределения (7.17) равна

$$\hat{\mathbf{a}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i, \qquad (7.19)$$

а оценки компонент вектора **b** могут быть найдены из системы уравнений, получающейся путем сравнения величин M_{ij} из (7.18) с соответствующими элементами выборочной корреляционной матрицы

$$\hat{\mathbf{M}} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{x}_i - \hat{\mathbf{a}})(\mathbf{x}_i - \hat{\mathbf{a}})'. \qquad (7.20)$$

Заметим, что число уравнений в такой системе будет, вообще говоря, больше числа неизвестных. Для решения может быть привлечен метод наименьших квадратов.

Определение оценки вектора **b** может быть упрощено, если воспользоваться тем обстоятельством, что этот вектор является наибольшим собственным вектором корреляцион-

ной матрицы распределения (7.17), причем для рассматриваемого случая сферической симметрии все характеристические числа корреляционной матрицы, за исключением наибольшего, равны между собой *).

Задача существенно упрощается, когда вектор среднего для одного из классов, например $\mathbf{a}_0$, известен (как в задаче обнаружения неизвестного сигнала на фоне шума, когда $\mathbf{a}_0 \equiv 0$). Тогда

$$\hat{\mathbf{b}} = \hat{\mathbf{a}} - \mathbf{a}_0 = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i - \mathbf{a}_0. \tag{7.21}$$

7.2.3. Неизвестные средние и корреляционные матрицы. Предположим сначала, что векторы средних двух p-мерных нормальных распределений известны и равны друг другу $\mathbf{a}_0 = \mathbf{a}_1 = \mathbf{a}$. Корреляционные матрицы этих распределений $\mathbf{M}_0$ и $\mathbf{M}_1$ неизвестны ($\mathbf{M}_0 \neq \mathbf{M}_1$). Имея классифицированную обучающую выборку $\mathbf{x}_1^{(0)}, \ldots, \mathbf{x}_{n_0}^{(1)}$ из первого распределения и $\mathbf{x}_1^{(1)}, \ldots, \mathbf{x}_{n_1}^{(1)}$ — из второго, можно записать оценки максимального правдоподобия неизвестных матриц $\mathbf{M}_0$ и $\mathbf{M}_1$ [см. (2.203)]:

$$\hat{\mathbf{M}}_0 = \frac{1}{n_0} \sum_{i=1}^{n_0} (\mathbf{x}_i^{(0)} - \mathbf{a}) (\mathbf{x}_i^{(0)} - \mathbf{a})', \tag{7.22}$$

$$\hat{\mathbf{M}}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} (\mathbf{x}_i^{(1)} - \mathbf{a}) (\mathbf{x}_i^{(1)} - \mathbf{a})'. \tag{7.22'}$$

Для того чтобы классифицировать наблюдение $\mathbf{X}$, можно воспользоваться оптимальным алгоритмом проверки гипотез о корреляционной матрице нормального распределения (см. задачу 1.7), заменив неизвестные корреляционные матрицы $\mathbf{M}_0$ и $\mathbf{M}_1$ их оценками (7.22) и (7.22′). Тогда получаем следующее правило классификации: наблюдение $\mathbf{X}$ относится ко второму распределению, если

$$(\mathbf{X} - \mathbf{a})' (\hat{\mathbf{M}}_1 - \hat{\mathbf{M}}_0) (\mathbf{X} - \mathbf{a}) \geqslant 2 \ln c + \ln \frac{\det \hat{\mathbf{M}}_0}{\det \hat{\mathbf{M}}_1}. \tag{7.23}$$

Совершая замену

$$\mathbf{Y} = \hat{\mathbf{f}}' (\mathbf{X} - \mathbf{a}), \quad \mathbf{Y}' = (y_1, \ldots, y_p),$$

*) О собственных векторах и собственных числах матриц см., например, в книге: А. Л. Мишиной и И. В. Проскурякова «Высшая алгебра». Физматгиз 1962, гл. II, § 1.

где матрица $\hat{\mathbf{f}}$ определяется из соотношения

$$\hat{\mathbf{M}}_1\hat{\mathbf{f}} = \hat{\mathbf{M}}_0\hat{\mathbf{f}}\hat{\mathbf{\Lambda}}$$

и $\hat{\mathbf{\Lambda}}$ — диагональная матрица, элементы $\hat{\lambda}_1, \ldots, \hat{\lambda}_p$ которой являются корнями уравнения

$$\det [\hat{\mathbf{M}}_1 - \hat{\lambda}\hat{\mathbf{M}}_0] = 0,$$

можно неравенство (7.23) привести к виду

$$\sum_{i=1}^{p} \left(1 - \frac{1}{\hat{\lambda}_i}\right) y_i^2 \geqslant 2 \ln c + \sum_{i=1}^{p} \ln \hat{\lambda}_i. \qquad (7.24)$$

Распределение статистики в левой части (7.24), а также вероятности ошибок, соответствующие правилу (7.24), изучались в [39]. Там, в частности, доказано, что при $n_0 \to \infty$ и $n_1 \to \infty$ правило (7.24) стремится по вероятности к правилу оптимальной проверки гипотез при известных $\mathbf{M}_0$ и $\mathbf{M}_1$ (см. задачу 1.7).

Если векторы средних двух нормальных распределений равны друг другу и неизвестны, то в (7.23) вместо a следует подставить его оценку по обучающим выборкам

$$\hat{\mathbf{a}} = \frac{n_0\hat{\mathbf{a}}_0 + n_1\hat{\mathbf{a}}_1}{n_0 + n_1}, \qquad (7.25)$$

где $\hat{\mathbf{a}}_0$ и $\hat{\mathbf{a}}_1$ определяются согласно (7.8).

В случае, когда векторы средних двух нормальных распределений не равны друг другу ($\mathbf{a}_0 \neq \mathbf{a}_1$) и неизвестны, а неизвестные корреляционные матрицы одинаковы ($\mathbf{M}_0 = \mathbf{M}_1 = \mathbf{M}$), используется классифицирующая статистика (7.9), в которой вместо $\mathbf{M}$ следует подставить ее оценку по обучающим выборкам, т. е. [см. (2.207)]

$$V = \left(\mathbf{X} - \frac{\hat{\mathbf{a}}_0 + \hat{\mathbf{a}}_1}{2}\right)' \hat{\mathbf{M}}^{-1} (\hat{\mathbf{a}}_1 - \hat{\mathbf{a}}_0), \qquad (7.26)$$

где

$$\hat{\mathbf{M}} = \frac{1}{n_0 + n_1 - 2} \times$$
$$\times \left[\sum_{i=1}^{n_0} (\mathbf{x}_i^{(0)} - \hat{\mathbf{a}}_0)(\mathbf{x}_i^{(0)} - \hat{\mathbf{a}}_0)' + \sum_{i=1}^{n_1} (\mathbf{x}_i^{(1)} - \hat{\mathbf{a}}_1)(\mathbf{x}_i^{(1)} - \hat{\mathbf{a}}_1)'\right] \qquad (7.27)$$

и $\hat{\mathbf{a}}_0$, $\hat{\mathbf{a}}_1$ определяются согласно (7.8). Распределение статистики (7.26) изучалось в [24, 34, 45].

7.2.4. Произвольное число распределений. Приведем соображения, позволяющие использовать обучение с учителем, в том случае, когда имеется $m+1$ классов $s_0, \ldots, s_m$, характеризуемых p-мерными распределениями (необязательно нормальными): $w_p(\mathbf{x} \mid s_0), \ldots, w_p(\mathbf{x} \mid s_m)$. Вид каждого из распределений известен, а все или часть их параметров неизвестны. Если получена классифицированная обучающая выборка, т. е. $m+1$ наборов p-мерных векторов, причем $\mathbf{x}_1^{(k)}, \ldots, \mathbf{x}_{n_k}^{(k)}$ принадлежит классу s_k, $k = 0, \ldots, m$, то можно по обучающим выборкам построить оценки неизвестных параметров распределений и, подставляя эти оценки в $w_p(\mathbf{x} \mid s_k)$, найти оценки функций распределения $\hat{w}_p(\mathbf{x} \mid s_k)$ (см. § 2.6.6). Далее, используя минимум среднего риска (при известных априорных вероятностях p_k, $k = 0, \ldots, m$, принадлежности каждому из распределений и величинах потерь Π_{ij}) как критерий качества классификации, можно, заменяя в байесовском правиле (1.144) функции распределения $w_p(\mathbf{x} \mid s_k)$ их оценками, получить следующее правило классификации: наблюдаемая выборка $\mathbf{X}$ принадлежит классу s_k, если

$$\sum_{i=0}^{m} (\Pi_{ij} - \Pi_{ik})\, p_i\, \hat{w}_p(\mathbf{X} \mid s_i) \geqslant 0, \qquad j = 0, \ldots, m; \quad j \neq k. \tag{7.28}$$

Пусть, например, имеется $m+1$ классов, характеризуемых многомерными нормальными распределениями с неизвестными средними $\mathbf{a}_k$ и корреляционными матрицами $\mathbf{M}_k = \mathbf{M}$ $(k = 0, \ldots, m)$, и пусть в результате обучения получены выборки из каждого распределения. Если выборка $\mathbf{x}_1^{(k)}, \ldots, \mathbf{x}_{n_k}^{(k)}$ принадлежит нормальному распределению с параметрами $\mathbf{a}_k$, $\mathbf{M}$ $(k = 0, \ldots, m)$, то можно построить оценки максимального правдоподобия неизвестных параметров:

$$\hat{\mathbf{a}}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} \mathbf{x}_i^{(k)}, \tag{7.29}$$

$$\hat{\mathbf{M}} = \frac{1}{\sum_{i=0}^{m} n_i - (m+1)} \sum_{k=0}^{m} \sum_{i=1}^{n_k} (\mathbf{x}_i^{(k)} - \hat{\mathbf{a}}_k)(\mathbf{x}_i^{(k)} - \hat{\mathbf{a}}_k)', \tag{7.30}$$

и найти оценки распределений [см. (2.201)]:

$$\hat{w}_p(\mathbf{x} \mid s_k) = \frac{1}{(2\pi)^{p/2} \sqrt{\det \hat{\mathbf{M}}}} \exp\left[-\frac{1}{2} (\mathbf{x} - \hat{\mathbf{a}}_k)' \hat{\mathbf{M}}^{-1} (\mathbf{x} - \hat{\mathbf{a}}_k) \right],$$
$$k = 0, \ldots, m. \tag{7.31}$$

Полагая $\Pi_{ij} = \Pi \, (i \neq j)$, $\Pi_{jj} = 0$, находим из (7.28) правило классификации: наблюдаемая выборка $\mathbf{X}$ принадлежит классу s_k, если

$$V_{kj} = \ln \frac{\hat{w}_p(\mathbf{x} \mid s_k)}{\hat{w}_p(\mathbf{x} \mid s_j)} = \left[\mathbf{x} - \frac{1}{2} (\hat{\mathbf{a}}_k + \hat{\mathbf{a}}_j) \right]' \hat{\mathbf{M}}^{-1} (\hat{\mathbf{a}}_k - \hat{\mathbf{a}}_j) \geqslant \ln \frac{p_j}{p_k},$$
$$j = 0, \ldots, m; \; j \neq k. \tag{7.32}$$

Когда все $n_k \to \infty$, совместное распределение статистик V_{kj} становится нормальным (см. [1, стр. 208]).

Если вид функций распределения $w_p(\mathbf{x} \mid s_k)$ неизвестен, то можно воспользоваться обучающими выборками для оценки этих функций по методике, указанной в § 2.7.4. Тогда в соответствии с (2.211)

$$\hat{w}_p(\mathbf{x} \mid s_k) = \frac{1}{n_k} \sum_{i=1}^{n_k} \prod_{j=1}^{p} \frac{1}{h(n)} K\left[\frac{x_j - x_{ij}^{(k)}}{h(n)} \right], \tag{7.33}$$

$$\mathbf{x} = (x_1, \ldots, x_p), \quad \mathbf{x}_i^{(k)} = (x_{i1}^{(k)}, \ldots, x_{ip}^{(k)}), \tag{7.34}$$

где ядро аппроксимации $K(y)$ и величина $h(n)$ удовлетворяют условиям (2.211′).

7.2.5. Связь с геометрическим подходом. Предположение о том, что классы характеризуются нормальными распределениями, имеет смысл и в тех случаях, когда в действительности эти распределения отличаются от нормального. Рассмотренные выше правила классификации, основанные на нормальных распределениях с неизвестными параметрами, являются оптимальными и для случаев, когда распределения на самом деле не нормальны, если критерием качества служит наименьшее «расстояние» от вектора наблюдений до вектора обучающих выборок [4]. Расстоянием вектора наблюдений $\mathbf{X}$ до вектора обучающей выборки $\mathbf{x}_1^{(i)}, \ldots$ $\ldots, \mathbf{x}_{n_i}^{(i)}$ называется величина

$$\frac{1}{n_i} \sum_{k=1}^{n_i} \| \mathbf{X} - \mathbf{x}_k^{(i)} \|^2,$$

(где $\| \mathbf{X} - \mathbf{x}_k^{(i)} \|$ — норма вектора $\mathbf{X} - \mathbf{x}_k^{(i)}$), т. е. сумма квадратов компонент этого вектора.

Естественно до классификации применить линейное преобразование для того, чтобы максимально уплотнить обучающие выборки данного класса. Точнее говоря, необходимо найти такое линейное преобразование $\mathbf{A}_i$ с сохранением объема (т. е. с якобианом преобразования, равным единице), при котором величина

$$r_i = \frac{1}{n_i(n_i-1)} \sum_{k=1}^{n_i} \sum_{j=1}^{n_i} \| \mathbf{A}_i (\mathbf{x}_k^{(i)} - \mathbf{x}_j^{(i)}) \|^2 \qquad (7.35)$$

минимальна. В [4] было доказано, что таким преобразованием является вращение $\mathbf{C}_i'$ с последующим диагональным преобразованием $\mathbf{D}_i$. Столбцами матрицы $\mathbf{C}_i$ служат собственные векторы *) выборочной корреляционной матрицы $\hat{\mathbf{M}}_i$ [см. (7.22)], а элементами диагональной матрицы $\mathbf{D}_i$ являются величины

$$d_{kj}^{(i)} = \frac{1}{\sigma_k^{(i)}} \left(\prod_{l=1}^{p} \sigma_l^{(i)} \right)^{\frac{1}{p}} \delta_{kj}, \qquad (7.36)$$

где $\sigma_l^{(i)}$ — среднеквадратическое отклонение векторов $\mathbf{x}_k^{(i)}$ в направлении l-го собственного вектора матрицы $\mathbf{C}_i$; p — размерность векторов и $\delta_{kj} = 1$ при $k = j$; $\delta_{kj} = 0$ при $k \neq j$.

Правило классификации может быть сформулировано теперь следующим образом: по обучающим выборкам и наблюдению $\mathbf{X}$ вычисляются величины

$$\rho_i = \frac{1}{n_i} \sum_{k=1}^{n_i} \| \mathbf{D}_i \mathbf{C}_i' (\mathbf{X} - \mathbf{x}_k^{(i)}) \|^2, \qquad i = 0, 1, \ldots, m, \qquad (7.37)$$

определяется наименьшее из них $\rho_j = \min_i \rho_i$; наблюдение относится к j-му классу.

В [4] доказано также, что указанное правило эквивалентно правилу, основанному на определении логарифма отношения правдоподобия, вычисленного в предположении нормальности распределений с выборочными векторами средних и корреляционных матриц (при равенстве априорных вероятностей $p_i = p = \frac{1}{m+1}$ и плат $\Pi_{ij} = \Pi$).

*) См. подстрочное примечание на стр. 454.

7.3. БАЙЕСОВСКИЕ ПРОЦЕДУРЫ ОБУЧЕНИЯ

7.3.1. Метод апостериорных вероятностей. Другой подход к проблеме классификации с учителем связан с допущением, что параметры $\theta_1, \ldots, \theta_n$, определяющие распределения классов $w_p(\mathbf{x} \mid \theta_1; s_1), \ldots, w_p(\mathbf{x} \mid \theta_m; s_m)$, представляют независимые случайные конечномерные векторы с известными априорными распределениями $w_M(\theta_k)$, $k = 1, \ldots, m$. Если имеется набор обучающих выборок $\mathbf{X}_{об} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$, где элемент $\mathbf{x}_k$ представляет вектор-строку $\mathbf{x}_k = (\mathbf{x}_1^{(k)}, \ldots, \mathbf{x}_{n_k}^{(k)})$, классифицированных выборок, принадлежащих классу s_k, и наблюдение $\mathbf{X}$, то, используя формулу Байеса, можно определить апостериорную вероятность класса s_k при заданных $\mathbf{X}_{об}$, $\mathbf{X}$:

$$P\{s_k \mid \mathbf{X}_{об}, \mathbf{X}\} = \frac{p_k W(\mathbf{X}_{об}, \mathbf{X} \mid s_k)}{\sum_{k=1}^{m} p_k W(\mathbf{X}_{об}, \mathbf{X} \mid s_k)}, \qquad (7.38)$$

где p_k — априорная вероятность принадлежности классу s_k. Имея величины $P\{s_k \mid \mathbf{X}_{об}, \mathbf{X}\}$, $k = 1, \ldots, m$, относим наблюдение $\mathbf{X}$ к тому классу s_j, для которого апостериорная вероятность максимальна, т. е.

$$P\{s_j \mid \mathbf{X}_{об}, \mathbf{X}\} > P\{s_k \mid \mathbf{X}_{об}, \mathbf{X}\} \qquad (7.39)$$

для всех $k \neq j$.

Так как

$$W(\mathbf{X}_{об}, \mathbf{X} \mid s_k) = W(\mathbf{X} \mid \mathbf{X}_{об}, s_k)\, W(\mathbf{X}_{об} \mid s_k), \qquad (7.40)$$

причем второй сомножитель в (7.40) зависит только от обучающих выборок, то из (7.38) и (7.39) следует, что алгоритм классификации сводится к вычислению величин $p_k W(\mathbf{X} \mid \mathbf{X}_{об}, s_k)$, $k = 1, \ldots, m$, и отнесению наблюдения $\mathbf{X}$ к тому классу s_j, которому соответствует максимальная из полученных величин. Если априорные вероятности p_k одинаковы, то классификация при заданном наборе $\mathbf{X}_{об}$ сводится к определению того класса s_j, для которого наблюдаемая выборка $\mathbf{X}$ максимизирует функцию правдоподобия $W(\mathbf{X} \mid \mathbf{X}_{об}, s_k)$. Последнюю можно рассматривать в качестве оценки неизвестных распределений классов при заданном наборе обучающих выборок $\mathbf{X}_{об}$.

Функцию $W(\mathbf{X} \mid \mathbf{X}_{об}, s_k)$ вычисляем, используя формулу полной вероятности:

$$W(\mathbf{X} \mid \mathbf{X}_{об}, s_k) = \int\limits_{\Omega_k} W(\mathbf{X} \mid \mathbf{X}_{об}, \theta_k, s_k)\, W(\theta_k \mid \mathbf{X}_{об}, s_k)\, d\theta_k =$$

$$= \int\limits_{\Omega_k} W(\mathbf{X} \mid \theta_k, s_k)\, W(\theta_k \mid \mathbf{x}_k, s_k)\, d\theta_k, \qquad \theta_k \in \Omega_k, \qquad (7.41)$$

так как очевидно, что $W(\theta_k \mid \mathbf{X}_{об}, s_k) = W(\theta_k \mid \mathbf{x}_k, s_k)$, а $W(\mathbf{X} \mid \mathbf{X}_{об}, \theta_k, s_k)$ вовсе не зависит от обучающих выборок.

Апостериорную плотность вероятности параметра θ_k для заданной обучающей выборки находим по формуле Байеса

$$W(\theta_k \mid \mathbf{x}_k, s_k) = w(\theta_k) \frac{W(\mathbf{x}_k \mid \theta_k, s_k)}{\int\limits_{\Omega_k} w(\theta_k)\, W(\mathbf{x}_k \mid \theta_k, s_k)\, d\theta_k}. \qquad (7.42)$$

В формуле (7.42) второй сомножитель указывает, каким образом в процессе обучения с учителем изменяется первоначальное априорное распределение параметров θ_k.

Заметим, что в том случае, когда параметры θ_k известны и равны θ_k^*, их условные плотности представляют дельта-функции

$$W(\theta_k \mid \mathbf{x}_k, s_k) = \delta(\theta_k - \theta_k^*)$$

и из (7.41) следует

$$W(\mathbf{X} \mid \mathbf{X}_{об}, s_k) = W(\mathbf{X} \mid \theta_k^*, s_k),$$

как и должно быть. В этом случае задача классификации полностью совпадает с многоальтернативной проверкой гипотез, рассмотренной кратко в первой главе.

7.3.2. Обнаружение неизвестного сигнала в нормальном шуме. В качестве простейшей *) иллюстрации указанной выше методики классификации рассмотрим задачу об обнаружении неизвестного сигнала в аддитивном нормальном некоррелированном шуме с нулевым средним и дисперсией σ^2. Сигнал представляет независимую от шума нормальную случайную величину a, априорное распределение которой определяется параметрами (a_0, σ_0^2). Проблема обнаружения сигнала состоит в этом случае в отнесении наблю-

---

*) Более общая задача распознавания неизвестных сигналов $(m \geqslant 2)$ в шумах рассмотрена в [2].

дения X к распределению смеси сигнала и шума $w_1(x \mid s_1)$ или к распределению шума $w_1(x \mid s_0)$.

Если априорное распределение сигнала не изменяется в процессе обучения, т. е.

$$W(a \mid x_1, \ldots, x_n; s_1) = \frac{1}{\sigma_0 \sqrt{2\pi}} e^{-\frac{(a-a_0)^2}{2\sigma_0^2}},$$

то в соответствии с (7.41) функции правдоподобия $W_1(X \mid s_1)$, $W_1(X \mid s_0)$ равны *)

$$W_1(X \mid s_1) = \frac{1}{2\pi\sigma_0\sigma} \int_{-\infty}^{\infty} \exp\left[-\frac{(a-a_0)^2}{2\sigma_0^2}\right] \exp\left[-\frac{(X-a)^2}{2\sigma^2}\right] da =$$

$$= \frac{1}{\sqrt{2\pi(\sigma_0^2+\sigma^2)}} \exp\left[-\frac{(X-a_0)^2}{2(\sigma_0^2+\sigma^2)}\right], \qquad (7.43)$$

$$W_1(X \mid s_0) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{X^2}{2\sigma^2}\right]. \qquad (7.43')$$

Правило выбора решения (классификации) по критерию максимума апостериорной вероятности является в этом случае частным случаем рассмотренных в гл. 1 и 5 байесовских правил проверки гипотез (обнаружения сигнала). Принимается решение о наличии сигнала, если

$$W_1(X \mid s_1) \geqslant W_1(X \mid s_0), \qquad (7.44)$$

и о том, что присутствует только шум, если выполняется неравенство, противоположное (7.44). Из (7.43), (7.43′) и (7.44) следует, что алгоритм классификации может быть записан следующим образом: наблюдение X относится к смеси сигнала и шума, если

$$a_0 X + \frac{1}{2}\frac{\sigma_0^2}{\sigma^2} X^2 \geqslant c, \qquad (7.45)$$

и оно относится к шуму, если

$$a_0 X + \frac{1}{2}\frac{\sigma_0^2}{\sigma^2} X^2 < c, \qquad (7.45')$$

где

$$c = a_0^2 + \frac{\sigma_0^2+\sigma^2}{2} \ln\left(1 + \frac{\sigma_0^2}{\sigma^2}\right). \qquad (7.45'')$$

---

*) Формула (7.43) следует непосредственно из того, что сумма сигнала и шума в рассматриваемом случае распределена нормально со средним, равным a_0, и дисперсией $\sigma_0^2 + \sigma^2$.

Если необходимо классифицировать не одно наблюдение, а выборку $(X_1, \ldots, X_N)$, то при сохранении предположений о независимых значениях сигнала и шума следует в (7.45) и (7.45′) заменить X и X^2 суммами наблюдаемых выборочных значений и квадратов этих значений [см. (1.127) и (5.182)]. Оптимальное устройство обнаружения сигнала состоит из двух блоков: согласованного фильтра и энергетического приемника [см. (5.87)]. Выходные значения суммируются, а сумма затем сравнивается с порогом.

Рассмотрим теперь ситуацию, когда априорные сведения о сигнале *уточняются в процессе обучения с учителем*. Пусть имеется выборка $x_1, \ldots, x_n$, о которой известно, что она принадлежит смеси сигнала и шума. Апостериорное распределение сигнала определяется по формуле (2.168)

$$W_1(a \mid x_1, \ldots, x_n; s_1) = \frac{\sqrt{n}}{\sigma\sqrt{2\pi}} \left(1 + \frac{\sigma^2}{n\sigma_0^2}\right)^{1/2} \times$$
$$\times \exp\left\{-\frac{n}{2\sigma^2}\left(1 + \frac{\sigma^2}{n\sigma_0^2}\right) \times \right.$$
$$\left. \times \left[a - \frac{1}{1 + \frac{\sigma^2}{n\sigma_0^2}} \left(\frac{1}{n}\sum_{k=1}^{n} x_k + \frac{a_0\sigma^2}{n\sigma_0^2}\right)\right]^2\right\}, \qquad (7.46)$$

т. е. так же, как и априорное распределение, является нормальным. Отличие состоит в том, что параметры апостериорного распределения зависят от обучающей выборки

$$a_n = m_1\{a\} = \frac{1}{1 + \frac{\sigma^2}{n\sigma_0^2}} \left(\frac{1}{n}\sum_{k=1}^{n} x_k + \frac{a_0\sigma^2}{n\sigma_0^2}\right), \qquad (7.47)$$

$$\sigma_n^2 = M_2\{a\} = \frac{\sigma^2}{n + \frac{\sigma^2}{\sigma_0^2}}. \qquad (7.48)$$

Отсюда следует, что функция правдоподобия $W_1(X \mid x_1, \ldots, x_n; s_1)$ — нормальная с параметрами a_n и σ_n^2 и что правило классификации наблюдения X при наличии обучающей выборки $x_1, \ldots, x_n$ получается из (7.45) и (7.45′) заменой a_0 величиной условного среднего a_n из (7.47) и σ_0^2 — величиной условной дисперсии σ_n^2 из (7.48). Иначе говоря, структура оптимального устройства обна-

ружения остается неизменной, а изменяются при обучении лишь параметры блоков, входящих в это устройство.

При $n \to \infty$ из (7.47) и (7.48) следует, что

$$a_n \sim \frac{1}{n} \sum_{i=1}^{n} x_i, \ \ \sigma_n^2 \to 0,$$

и, следовательно, при неограниченном увеличении размера обучающей выборки оптимальное устройство обнаружения приближается к линейному согласованному фильтру (второе слагаемое в левой части (7.45) будет стремиться к нулю).

Рассмотренный пример можно обобщить на случай коррелированных значений сигнала и коррелированного шума при сохранении предположения об аддитивности шума и независимости его от сигнала. Пусть вектор средних значений шума нулевой, а его корреляционная матрица равна $\mathbf{M}$. Априорное многомерное нормальное распределение сигнала характеризуется вектором средних $\mathbf{a}_0$ и корреляционной матрицей $\mathbf{M}_0$. Априорное распределение аддитивной смеси независимых сигнала и шума также нормальное с вектором средних $\mathbf{a}_0$ и корреляционной матрицей $\mathbf{M} + \mathbf{M}_0$. Имеется векторная обучающая выборка $\mathbf{x}_1, \ldots, \mathbf{x}_n$, принадлежащая смеси сигнала и шума. Тогда апостериорное распределение сигнала после обучения также нормальное, но с измененными (условными) вектором средних и корреляционной матрицей. Они определяются из следующего рекуррентного соотношения, справедливого для условных вектора средних $\mathbf{a}_n$ и корреляционной матрицы $\mathbf{M}_n$ нормального распределения [1]:

$$\mathbf{a}_n = \mathbf{M}(\mathbf{M}_{n-1} + \mathbf{M})^{-1}\mathbf{a}_{n-1} + \mathbf{M}_{n-1}(\mathbf{M}_{n-1} + \mathbf{M})^{-1}\mathbf{x}_n, \quad (7.49)$$

$$\mathbf{M}_n = \mathbf{M}(\mathbf{M}_{n-1} + \mathbf{M})^{-1}\mathbf{M}_{n-1}. \quad (7.50)$$

Из (7.49) и (7.50) следует:

$$\mathbf{a}_n = \frac{\mathbf{M}}{n}\left(\mathbf{M}_0 + \frac{\mathbf{M}}{n}\right)^{-1} + \mathbf{M}_0\left(\mathbf{M}_0 + \frac{\mathbf{M}}{n}\right)^{-1}\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i, \quad (7.51)$$

$$\mathbf{M}_n = \frac{\mathbf{M}}{n}\left(\mathbf{M}_0 + \frac{\mathbf{M}}{n}\right)^{-1}\mathbf{M}_0. \quad (7.52)$$

Особенно простой вид формулы (7.51) и (7.52) приобретают в том случае, когда $\mathbf{M} = \lambda\mathbf{M}_0$, где λ — положительное

число. Тогда

$$\mathbf{a}_n = \frac{\lambda}{\lambda+n}\,\mathbf{a}_0 + \frac{n}{\lambda+n}\cdot\frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i, \qquad (7.53)$$

$$\mathbf{M}_n = \frac{1}{\lambda+n}\,\mathbf{M}. \qquad (7.54)$$

В одномерном случае (7.53) переходит в (7.47), а (7.54) — в (7.48), причем $\lambda = \sigma^2/\sigma_0^2$.

Функция правдоподобия $W(\mathbf{X}\,|\,\mathbf{X}_{об};\, s_1)$ в общем случае является многомерной нормальной функцией распределения, характеризуемой вектором средних $\mathbf{a}_n$ и корреляционной функцией $\mathbf{M}_n$, определяемыми по формулам (7.51) и (7.52). Следует подчеркнуть, что только условные средние зависят от обучающих выборок (точнее, от выборочного среднего), а условная корреляционная матрица зависит лишь от размера n обучающей выборки.

7.3.3. Простой перебор и метод угадывания. Рассмотрим теперь задачу обнаружения неизвестного сигнала в шумах при обучении без учителя, когда обучающая выборка *не классифицирована*. Это означает, что заранее неизвестно, принадлежит ли данное выборочное значение x_i смеси сигнала и шума или только шуму. При этом на пути построения правила классификации наблюдения X по критерию максимальной апостериорной вероятности сразу же возникает трудность в определении апостериорной плотности параметра $W_1(a\,|\,\mathbf{X}_{об};\, s_1)$ для неклассифицированной выборки $\mathbf{X}_{об} = (x_1, \ldots, x_n)$.

Самый простой выход из этого затруднения состоит, как кажется, в переборе всех возможных последовательностей $x_1, \ldots, x_n$, любой член x_i которой может принадлежать либо смеси сигнала и шума (состояние s_1), либо только шуму (состояние s_0). Для любой частной последовательности $\Pi_n^{(k)}$ состояний при обучении

$$\Pi_n^{(k)} = (s^{(1)}, \ldots, s^{(n)}),$$

где каждый из элементов $s^{(j)}$, $j = 1, \ldots, n$ может совпадать либо с s_1, либо с s_0, апостериорная плотность $W(a\,|\,\mathbf{X}_{об};\, \Pi_n^{(k)})$ вычисляется по методике, указанной в § 7.3.2, т. е. как при обучении с учителем. Число различных последовательностей для обучающей выборки раз-

мером n равно 2^n. Тогда

$$W_1(a \mid \mathbf{X}_{об}) = \sum_{k=1}^{2^n} W_1(a \mid \mathbf{X}_{об}; \Pi_n^{(k)}) P\{\Pi_n^{(k)} \mid \mathbf{X}_{об}\}, \qquad (7.55)$$

где $P\{\Pi_n^{(k)} \mid \mathbf{X}_{об}\}$ — вероятность того, что в обучающей выборке реализуется последовательность состояний $\Pi_n^{(k)}$. (Эта вероятность может быть легко вычислена, если пред-

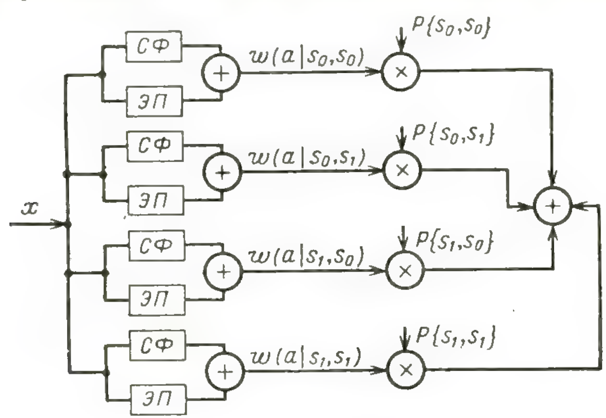


Рис. 31. Блок-схема адаптивного устройства обнаружения неизвестного сигнала при обучении без учителя.

положить, например, что появление или непоявление сигнала при обучении независимо и известна априорная вероятность появления сигнала в любом данном наблюдении.)

Из (7.55) следует, что оптимальное устройство обнаружения при самообучении на выборке размером n должно содержать 2^n устройств, каждое из которых такое же, как и при обучении с учителем (согласованный фильтр и энергетический приемник). Выходы указанных устройств умножаются на весовые коэффициенты $P\{\Pi_n^{k)} \mid \mathbf{X}_{об}\}$ и затем суммируются. Блок-схема такого устройства для простого случая обучающей выборки из двух элементов показана на рис. 31. При $n = 3$ устройство стало бы примерно вдвое сложнее. Для размера обучающей выборки порядка нескольких десятков реализуемость устройства, основанного на переборе, становится по меньшей мере проблематичной.

Естественно, что при малом размере n обучающей выборки рабочие характеристики классификации оказываются плохими, а по мере увеличения n они приближаются к характеристикам алгоритмов с полностью известной статистикой распознаваемых классов. При этом, однако, сложность оптимального устройства при обучении без учителя растет как 2^n.

Было показано [44], что для определения апостериорной плотности неизвестного параметра класса, можно указать конечный вычислительный алгоритм классификации, не зависящий от размера обучающей выборки, тогда и только тогда, когда обучающие наблюдения $x_1, \ldots, x_n$ характеризуются достаточной статистикой конечной размерности. Для рассматриваемой здесь задачи обнаружения неизвестного сигнала плотность вероятности любого элемента обучающей выборки (без учителя) имеет вид [см. (7.13)]

$$w_1(x_i \mid a) = \frac{p}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - a)^2}{2\sigma^2}\right] + \frac{1-p}{\sigma\sqrt{2\pi}} \exp\left[-\frac{x_i^2}{2\sigma^2}\right],$$
$$0 < p < 1, \tag{7.56}$$

где p — вероятность наличия сигнала. Но, как доказано в работе [11], для плотностей вероятностей вида (7.56) не существует достаточной статистики конечной размерности. Отсюда следует, что сложность любого оптимального устройства обнаружения при обучении без учителя должна неограниченно возрастать при увеличении размера обучающей выборки.

Рассмотрим одну из возможностей упрощения правила классификации при обучении без учителя ценой потери оптимальности этого правила. Так как оптимальное устройство обнаружения неизвестного сигнала при обучении с учителем имеет приемлемо простую структру, а аналогичное устройство при самообучении необозримо усложняется, то в последнем случае естественной представляется процедура, состоящая в угадывании последовательности состояний (s_1 — смесь сигнала и шума, s_0 — только шум) при переходе от одного элемента обучающей выборки к другому и в использовании этой отгадки так, как будто бы она правильна (как если бы она была «учителем») *). Структура

*) По этой причине метод угадывания иногда называют обучением с реальным учителем, т. е. с таким учителем, который может ошибаться.

устройства обнаружения становится при этом тождественной структуре оптимального устройства обнаружения с учителем, причем результаты отгадок при самообучении используются для изменения параметров этого устройства по формулам, указанным в § 7.3.2. Но так как при угадывании неизбежны ошибки, устройство обнаружения теряет свойство оптимальности. Однако, как отмечено в [44], для удовлетворительной работы такого устройства обнаружения нет необходимости, чтобы угаданная последовательность состояний была в точности правильной. Было доказано, что при больших размерах n обучающей выборки существует подмножество множества 2^n всех возможных последовательностей состояний, обладающее двумя свойствами: 1) вероятность присутствия истинной последовательности состояний в этом подмножестве близка к единице, 2) если отгаданная последовательность является элементом этого подмножества, то устройство обнаружения, использующее метод отгадывания, при $n \to \infty$ стремится к оптимальному. К сожалению, пока не найдены эффективные методы отыскания подмножеств, обладающих указанными свойствами.

Видоизменение метода угадывания последовательности состояний в обучающей выборке представляет подход, который базируется на предположении, что последовательность *решений* в процессе самообучения является серией догадок [41]. При таком подходе параметрам устройства обнаружения неизвестного сигнала вначале приписываются произвольные значения. Каждый раз, когда устройство выдает решение о наличии сигнала, параметры устройства изменяются в соответствии с принятым в этот момент сигналом. Аналогичный подход с небольшими изменениями использовался в [29].

7.3.4. Адаптивный байесовский подход. Обычный байесовский подход, который использовался во всех предыдущих главах книги, требует знания априорных распределений неизвестных параметров. В § 7.3.1 этой главы начальное априорное распределение параметра класса видоизменялось при обучении с учителем. Нельзя ли в задачах классификации воспользоваться обучением с учителем для того, чтобы использовать байесовский подход без какого-либо начального априорного распределения? Утвердительный ответ на поставленный вопрос базируется на одной теореме, установленной С. Н. Бернштейном и Р. Мизесом [14].

Сущность ее состоит в следующем. Пусть x_1 — выборочное значение из распределения с неизвестным (случайным) параметром θ. Апостериорное распределение этого параметра равно

$$W_1(\theta \mid x_1) = w_1(\theta) \frac{W_1(x_1 \mid \theta)}{\int_{-\infty}^{\infty} w_1(\theta) W_1(x_1 \mid \theta)\, d\theta}, \qquad (7.57)$$

где $w_1(\theta)$ — априорная плотность вероятности параметра θ. Если извлекается следующее выборочное значение x_2, то $W_1(\theta \mid x_1)$ может быть использовано в качестве нового априорного распределения для вычисления апостериорного

$$W_1(\theta \mid x_1, x_2) = W_1(\theta \mid x_1) \frac{W_1(x_2 \mid \theta)}{\int_{-\infty}^{\infty} W_1(\theta \mid x_1) W_1(x_2 \mid \theta)\, d\theta} =$$

$$= w_1(\theta) \frac{W_1(x_1 \mid \theta) W_1(x_2 \mid \theta)}{\int_{-\infty}^{\infty} w_1(\theta) W_1(x_1 \mid \theta) W_1(x_2 \mid \theta)\, d\theta}. \qquad (7.58)$$

Аналогично, когда имеется выборка n независимых элементов $x_1, \ldots, x_n$,

$$W_1(\theta \mid x_1, \ldots, x_n) = w_1(\theta) \frac{W_n(x_1, \ldots, x_n \mid \theta)}{\int_{-\infty}^{\infty} w_1(\theta) W_n(x_1, \ldots, x_n \mid \theta)\, d\theta}, \qquad (7.59)$$

где

$$W_n(x_1, \ldots, x_n \mid \theta) = \prod_{k=1}^{n} W_1(x_i \mid \theta).$$

Упомянутая выше теорема утверждает, что если априорная плотность $W_1(\theta)$ параметра θ непрерывна, то по мере возрастания объема выборки апостериорное распределение $W_1(\theta \mid x_1, \ldots, x_n)$ перестает зависеть от априорного распределения. Таким образом, если только существует непрерывная априорная плотность параметра, то при достаточно большом n более или менее безразлично, какую функцию $w_1(\theta)$ подставить в формулу (7.59). Эта предельная теорема послужила, по-видимому, идейной основой адаптивного байесовского подхода *) к проблеме классификации, предложенного Роббинсом [14, 16].

*) Часто рассматриваемый метод называют эмпирическим байесовским подходом.

Пусть $(x_1, \theta_1), \ldots, (x_n, \theta_n)$ — последовательность независимых пар случайных величин, причем все θ_i, $i = 1, \ldots, n$, подчиняются одному и тому же априорному (неизвестному) распределению $w_1(\theta)$, а все x_i — распределению

$$W_1(x) = \int_{-\infty}^{\infty} W_1(x \mid \theta) w_1(\theta) d\theta.$$

Для дискретных x_i

$$P\{x_i = x\} = \int_{-\infty}^{\infty} P\{x_i = x \mid \theta\} w_1(\theta) d\theta.$$

Если принимается решение γ о том, что наблюдение X принадлежит распределению $W_1(x \mid \theta)$, то возникают потери, определяемые неотрицательной функцией $\Pi(\gamma, \theta)$. Так как $w_1(\theta)$ неизвестно, то правило выбора решения строится по обучающей выборке $x_1, \ldots, x_n$ (величины $\theta_1, \ldots, \theta_n$ остаются всегда неизвестными). Желательно, чтобы это правило при неограниченном увеличении размера обучающей выборки приближалось к байесовскому правилу выбора решения, когда $w_1(\theta)$ известно. Правило, удовлетворяющее этому требованию, назовем асимптотически оптимальным. В работе [17] указан способ построения асимптотически оптимальных правил. Ограничимся изложением этого способа применительно к двухальтернативной задаче теории статистических решений с неизвестным априорным распределением параметра θ, определяющего статистику наблюдений (работа [17] содержит общую постановку).

Обозначим

$$\Delta(x) = \int_{-\infty}^{\infty} [\Pi(\gamma_1, \theta) - \Pi(\gamma_0, \theta)] W_1(x \mid \theta) w_1(\theta) d\theta \qquad (7.60)$$

и предположим, что функция $\Delta_n(X)$ от наблюдения X, вид которой зависит от обучающей выборки $x_1, \ldots, x_n$, сходится по вероятности к $\Delta(X)$ при $n \to \infty$. Тогда асимптотически оптимальным правилом выбора решения будет следующее: принимается решение γ_0, если

$$\Delta_n(X) \geqslant 0, \qquad (7.61)$$

и принимается решение γ_1, если выполняется неравенство, противоположное (7.61). (Для дискретных распределений $W_1(x \mid \theta)$ в (7.60) следует заменить на $P\{x_i = x \mid \theta\}$.)

Рассмотрим пример, иллюстрирующий возможность фактического построения последовательности функций $\Delta_n(x)$, сходящихся по вероятности к $\Delta(x)$. Проверяется односторонняя гипотеза H_0 о том, что параметр θ пуассоновского распределения дискретной случайной величины x не превосходит заданной величины θ_0, т. е. $\theta \leqslant \theta_0$. Пусть γ_0 — решение, состоящее в принятии гипотезы H_0, а γ_1 — в отклонении этой гипотезы. Зададим функцию потерь $\Pi(\gamma_i, \theta)$ в виде

$$\Pi(\gamma_0, \theta) = \begin{cases} 0, & \theta \leqslant \theta_0, \\ \theta - \theta_0, & \theta \geqslant \theta_0, \end{cases} \tag{7.62}$$

$$\Pi(\gamma_1, \theta) = \begin{cases} \theta_0 - \theta, & \theta \leqslant \theta_0, \\ 0, & \theta \geqslant \theta_0. \end{cases} \tag{7.62'}$$

Так как для пуассоновского распределения дискретной случайной величины

$$P\{x = x_i \mid \theta\} = \frac{\theta^x}{x!} e^{-\theta}, \ \theta > 0, \qquad x = 0,\ 1,\ 2,\ \ldots, \tag{7.63}$$

то, подставляя (7.63) в (7.60) [вместо $W_1(x \mid \theta)$], получим с учетом (7.62), (7.62′)

$$\Delta(x) = \int_0^\infty (\theta_0 - \theta) \frac{\theta^x}{x!} e^{-\theta} w_1(\theta)\, d\theta = \theta_0 F_1(x) - (x+1) F_1(x+1), \tag{7.64}$$

где

$$F_1(x) = \int_0^\infty \frac{\theta^x}{x!} e^{-\theta} w_1(\theta)\, d\theta = P\{x_i = x\}. \tag{7.65}$$

Введем счетчик совпадений

$$\nu(x, y) = \begin{cases} 1, & x = y \\ 0, & x \neq y \end{cases} \tag{7.66}$$

и рассмотрим следующую функцию от x, зависящую от обучающей выборки $x_1, \ldots, x_n$:

$$u_n(x) = \frac{1}{n} \sum_{i=1}^{n} \nu(x, x_i). \tag{7.67}$$

Сумма в правой части (7.67) равна числу тех обучающих выборочных значений, которые в точности равны x, и, следовательно, $u_n(x)$ представляет эмпирическое распределение x, сходящееся при $n \to \infty$ по вероятности к $F_1(x)$ [см. (7.65)]. Тогда функция

$$\Delta_n(x) = \theta_0 u_n(x) - (x+1) u_n(x+1) \qquad (7.68)$$

сходится по вероятности к $\Delta(x)$, определенной согласно (7.64). Отсюда следует асимптотически оптимальное правило: при наличии обучающей выборки $x_1, \ldots, x_n$ принимается решение γ_0 (справедлива гипотеза H_0), если

$$\frac{\theta_0}{n} \sum_{i=1}^{n} \nu(X, x_i) - \frac{X+1}{n} \sum_{i=1}^{n} \nu(X+1, x_i) \geqslant 0, \qquad (7.69)$$

и решение γ_1 (гипотеза H_0 неверна), если имеет место неравенство, противоположное (7.69).

7.4. СТОХАСТИЧЕСКАЯ АППРОКСИМАЦИЯ

7.4.1. Нули и экстремумы функции регрессии. Пусть каждому значению параметра θ соответствует случайная величина x с функцией распределения $F_1(x \mid \theta)$, причем $m_1\{x \mid \theta\} = m(\theta)$. Требуется найти решение уравнения регрессии

$$m(\theta) = 0 \qquad (7.70)$$

в предположении, что оно имеет единственный корень и что функции $F_1(x \mid \theta)$ и $m(\theta)$ неизвестны. Роббинс и Монро предложили [13] итерационную процедуру так называемой *стохастической аппроксимации* для построения оценок указанного корня при помощи обучающей выборки $x_1, \ldots, x_n$, для каждого элемента x_k которой

$$m_1\{x_k \mid \hat{\theta}_k\} = m(\hat{\theta}_k).$$

Оценка $\hat{\theta}_{n+1}$ искомого корня находится из оценки $\hat{\theta}_n$ по обучающему выборочному значению x_n из соотношения

$$\hat{\theta}_{n+1} = \hat{\theta}_n + a_n x_n, \quad n \geqslant 1, \qquad (7.71)$$

причем $\hat{\theta}_1 = \theta_1$ — произвольное постоянное число. Если коэффициенты a_n удовлетворяют условиям

$$\sum_{n=1}^{\infty} a_n = \infty, \quad \sum_{n=1}^{\infty} a_n^2 < \infty, \qquad (7.72)$$

то при неограниченном увеличении размера n обучающей выборки оценка $\hat{\theta}_{n+1}$ стремится по вероятности к корню уравнения (7.70). Заметим, что условиям (7.72) удовлетворяет, например, гармонический ряд $\left(a_n = \frac{1}{n}\right)$.

Аналогичная итерационная процедура может быть использована для определения экстремума унимодальной функции регрессии $m(\theta)$. Такая процедура, предложенная Кифером и Вольфовицем [13], позволяет определить оценку $\hat{\theta}_{n+1}$ экстремального значения функции регрессии в зависимости от предыдущей оценки $\hat{\theta}_n$ и обучающих выборочных значений x_{2n} и x_{2n-1} из соотношения

$$\hat{\theta}_{n+1} = \hat{\theta}_n + a_n \frac{x_{2n} - x_{2n-1}}{c_n}, \qquad n \geqslant 1, \tag{7.73}$$

причем выборки x_{2n} и x_{2n-1} независимы и соответствуют значениям параметра $\hat{\theta}_n + c_n$ и $\hat{\theta}_n - c_n$. Начальная величина θ_1, как и в предыдущем случае, произвольная. Если коэффициенты a_n и c_n удовлетворяют условиям

$$a_n > 0, \quad c_n > 0, \quad c_n \to 0 \quad \text{при} \quad n \to \infty, \tag{7.74}$$

$$\sum_{n=1}^{\infty} a_n = \infty, \quad \sum_{n=1}^{\infty} a_n c_n < \infty, \quad \sum_{n=1}^{\infty} \left(\frac{a_n}{c_n}\right)^2 < \infty, \tag{7.74'}$$

то при неограниченном увеличении размера обучающей выборки оценка $\hat{\theta}_{n+1}$, определенная согласно (7.73), стремится по вероятности к экстремальному значению функции регрессии.

При некоторых дополнительных требованиях к функции $m(\theta)$ оценки $\hat{\theta}_{i+1}$, определяемые согласно (7.71) и (7.73), могут сходиться в среднеквадратическом к нулю и экстремуму функции регресссии соответственно. Имеются также обобщения указанных процедур на многомерный случай (см. [13]).

Рассмотренные алгоритмы могут быть использованы и для оценки нуля или экстремума функции $m_1\{f(x) \mid \theta\}$, где $f(x)$ — заданная функция случайной величины x с распределением $F_1(x \mid \theta)$. В этом случае в формулах (7.71) и (7.73) вместо обучающих выборочных значений x_k следует подставить $f(x_k)$.

Недостатком методов стохастической аппроксимации является пока отсутствие каких-либо систематических ука-

заний на способы выбора коэффициентов a_n и c_n в (7.71) и (7.73), обеспечивающих достаточно быструю сходимость итерационных процессов, а также способов вычисления точности оценок, которые позволяли бы на каком-то шаге остановить этот процесс. Однако достоинством рассмотренных алгоритмов оценки является их простота, позволяющая относительно легко реализовать итерационный процесс.

7.4.2. Оценка параметров и функций распределения. Стохастическая аппроксимация может быть использована для оценки неизвестных параметров и функций распределения классов [20]. В качестве простейшего примера рассмотрим сначала оценку неизвестного среднего значения. Пусть параметр в (7.70) представляет неизвестное среднее. Тогда, полагая $f(x) = x - \theta_0$, получим уравнение регрессии

$$m_1\{x - \theta_0 \mid \theta\} = 0,$$

и из (7.71) находим рекуррентное соотношение для оценки неизвестного среднего (корня уравнения регрессии):

$$\hat{\theta}_{n+1} = \hat{\theta}_n + a_n(x_n - \hat{\theta}_n) = \hat{\theta}_n(1 - a_n) + a_n x_n. \qquad (7.75)$$

При $a_n = \frac{1}{n}$ из (7.75) следует, что

$$\hat{\theta}_{n+1} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad (7.75')$$

т. е. оценка среднего значения равна среднему арифметическому обучающих выборочных значений.

Если σ^2 — неизвестная дисперсия случайной величины x с нулевым средним, то, полагая $f(x) = x^2 - \sigma_0^2$, получим уравнение регрессии

$$m_1\{x^2 - \sigma_0^2 \mid \sigma^2\} = 0$$

и соответствующий ему алгоритм оценки дисперсии

$$\hat{\sigma}_{n+1}^2 = \hat{\sigma}_n^2 + a_n(x_n^2 - \hat{\sigma}_n^2) = \hat{\sigma}_n^2(1 - a_n) + a_n x_n^2. \qquad (7.76)$$

При $a_n = \frac{1}{n}$ из (7.76) следует:

$$\hat{\sigma}_{n+1}^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2. \qquad (7.76')$$

Можно применить стохастическую аппроксимацию и для оценки неизвестной плотности $w_1(x)$ по обучающим выбор-

кам, принадлежащим этой плотности, если воспользоваться разложением в ряд по ортогональным функциям $Q_k(x)$ [см. (2.120) в первой книге]. Тогда оценка неизвестной плотности сводится к оценке коэффициентов разложения

$$c_n^{(k)} = m_1\{Q_k(x)\} \tag{7.77}$$

из соотношения

$$\hat{c}_{n+1}^{(k)} = \hat{c}_n^{(k)} + a_n^{(k)}[Q_k(x_n) - \hat{c}_n^{(k)}], \qquad k = 1,\ 2,\ \ldots \tag{7.78}$$

При $a_n^{(k)} = \frac{1}{n}$

$$\hat{c}_{n+1}^{(k)} = \frac{1}{n}\sum_{i=1}^{n} Q_k(x_i). \tag{7.78'}$$

7.4.3. Различение двух неизвестных сигналов. Для иллюстрации итерационной процедуры (7.73) рассмотрим задачу о различении двух сигналов s_0 и s_1, функции распределения которых $w_1(x|s_0)$ и $w_1(x|s_1)$ неизвестны. Правило выбора решения состоит в сравнении наблюдения X с порогом θ. Если $X < \theta$, то наблюдение относится к сигналу s_0, а если $X \geqslant \theta$, то — к сигналу s_1. Необходимо, используя обучающие выборки, выбрать порог θ таким образом, чтобы минимизировать средний риск.

Согласно (1.18) средний риск в рассматриваемом случае равен

$$R(\theta) = \int_{\theta}^{\infty} [q\Pi_{01}w_1(x|s_0) + p\Pi_{11}w_1(x|s_1)]\,dx + \int_{-\infty}^{\theta} [q\Pi_{00}w_1(x|s_0) + p\Pi_{10}w_1(x|s_1)]\,dx, \tag{7.79}$$

где p и q — априорные вероятности появления сигналов s_1 и s_0, а Π_{ij}, $(i,\ j = 0,\ 1)$ — элементы платежной матрицы.

Вводя решающую функцию (см. § 1.2.1)

$$\Phi(x, \theta) = \begin{cases} 0, & x < \theta, \\ 1, & x \geqslant \theta, \end{cases} \tag{7.80}$$

можно выражение среднего риска записать в виде

$$R(\theta) = \int_{-\infty}^{\infty} \{\Phi(x, \theta)[q\Pi_{01}w_1(x|s_0) + p\Pi_{11}w_1(x|s_1)] + [1 - \Phi(x, \theta)][q\Pi_{00}w_1(x|s_0) + p\Pi_{10}w_1(x|s_1)]\}\,dx. \tag{7.81}$$

Из (7.81) следует, что средний риск $R(\theta)$ представляет функцию регрессии для дискретной случайной величины с четырьмя возможными значениями

$$y = f(x) = \Pi_{ij}, \quad i, j = 0, 1, \tag{7.82}$$

когда x — выборочное значение из распределения сигнала s_i и

$$\Phi(x, \theta) = j. \tag{7.82'}$$

Таким образом,

$$R(\theta) = m_1\{f(x) \mid \theta\}, \tag{7.83}$$

и задача об определении оптимального порога θ^* для классификации наблюдения сводится к вычислению минимума функции регрессии (7.83) при неизвестных распределениях сигналов.

Оставляя в стороне исследование функции регрессии, имеющей несколько экстремумов, ограничимся случаем единственного минимума, как показано на рис. 32

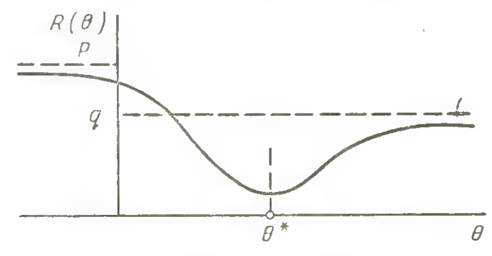


Рис. 32. Функция среднего риска.

(при $\Pi_{10} = \Pi_{01} = 1$ и $\Pi_{00} = \Pi_{11} = 0$). Для оценки оптимального порога θ^* может быть использован итерационный алгоритм (7.73)

$$\hat{\theta}_{n+1} = \hat{\theta}_n + a_n \frac{y_{2n} - y_{2n-1}}{c_n}, \tag{7.84}$$

где $y_{2n} = \Pi_{ij}$, когда x_{2n} — выборочное значение из сигнала s_i и $\Phi(x_{2n}, \hat{\theta}_n + c_n) = j$, $y_{2n-1} = \Pi_{ij}$, когда x_{2n-1} — выборочное значение из сигнала s_i и $\Phi(x_{2n-1}, \hat{\theta}_n - c_n) = j$.

Оценка $\hat{\theta}_1$ на первом шаге представляет произвольную постоянную, а коэффициенты a_n и c_n подчинены условиям (7.74), (7.74′).

При неограниченном увеличении размера n обучающих выборок оценка $\hat{\theta}_{n+1}$ стремится по вероятности к оптималь-

ному порогу θ^*, для которого средний риск $R(\theta^*)$ имеет минимально возможную величину. При дополнительных предположениях о функции $R(\theta)$ сходимость может обеспечиваться и в среднеквадратическом (см. [26]).

ЛИТЕРАТУРА

МОНОГРАФИИ

1. Андерсон Т. Введение в многомерный статистический анализ. Пер. с англ., под ред. Б. В. Гнеденко. Физматгиз, 1961.
2. Браверман Д. Теория опознавания образов. В сб. «Статистическая теория связи». Пер. с англ., под ред. А. В. Балакришнана. Изд-во «Мир», 1967.
3. Куликовский Р. Оптимальные и адаптивные процессы в системах автоматического регулирования. Пер. с польского, под ред. А. Г. Бутковского. Изд-во «Наука», 1967.
4. Себастьян Г. С. Процессы принятия решений при распознавании образов. Пер. с англ., под ред. В. И. Иваненко. Изд-во «Техника», Киев, 1965.

СТАТЬИ

5. Айзерман М. А., Браверман Э. М., Розоноэр Л. И. Теоретические основы метода потенциальных функций в задаче обучения автоматов разделению входных ситуаций на классы. «Автоматика и телемеханика», 1964, № 6.
6. Айзерман М. А., Браверман Э. М., Глушков В. М., Ковалевский В. А., Летичевский А. А. Теория опознавания образов и обучающихся систем. «Известия АН СССР», Техническая кибернетика, 1963, № 5.
7. Айзерман М. А., Браверман Э. М., Розоноэр Л. И. Метод потенциальных функций в задаче о восстановлении характеристики функционального преобразователя по случайно наблюдаемым точкам. «Автоматика и телемеханика», 1964, № 12.
8. Айзерман М. А., Браверман Э. М., Розоноэр Л. И. Процесс Роббинса — Монро и метод потенциальных функций. «Автоматика и телемеханика», 1965, № 11.
9. Башаринов А. Е. Об асимптотических экстремальных процедурах параметрического распознавания образов. «Радиотехника и электроника», 1965, № 5.
10. Бернштейн С. Н. О доверительных вероятностях Фишера. «Известия АН СССР», сер. математическая, 1941, № 5.
11. Дынкин Е. Б. Необходимые и достаточные статистики для семейства распределений вероятностей. «Успехи математических наук», 1951.
12. Лихарев В. А., Тебякин В. П. Методы обучения систем, решающих задачи распознавания образов и обнаружения сигналов на фоне шумов (обзор). «Зарубежная радиоэлектроника», 1967, № 11.

13. Логинов Н. В. Методы стохастической аппроксимации. «Автоматика и телемеханика», 1966, № 4.
14. Нейман Дж. Два прорыва в теории выбора статистических решений. Пер. с англ. «Математика», 1964, 8 : 2.
15. Пугачев В. С. Оптимальные алгоритмы обучения автоматических систем в случае неидеального учителя. Докл. АН СССР, 1967, т. 172, № 5.
16. Роббинс Г. Эмпирический байесовский подход к статистике. Пер. с англ. «Математика», 1964, 8 : 2.
17. Роббинс Г. Эмпирический байесовский подход к задачам теории статистических решений. Пер. с англ. «Математика», 1966, 10 : 5.
18. Харкевич А. А. О выборе признаков при машинном опознавании. «Техническая кибернетика», 1963, № 2.
19. Цыпкин Я. З. О восстановлении характеристики функционального преобразователя по случайно наблюдаемым точкам. «Автоматика и телемеханика», 1965, № 11.
20. Цыпкин Я. З. Адаптация, обучение и самообучение в автоматических системах. «Автоматика и телемеханика», 1966, № 1.
21. Чжоу С. К. Оптимальная система распознавания знаков с помощью решающих функций. Пер. с англ. «Математика», 1959, 3 : 1.
22. Abramson N., Braverman D. Learning to recognize patterns in a random environment. IRE Trans., 1962, IT-8, № 5.
23. Abramson N., Braverman D., Sebastyen G. Pattern recognition and machine learning. IEEE Trans., 1963, IT-9, № 4.
24. Anderson T. W. Classification by multivariate analysis. Psychometrika, 1951, v. 16.
25. Anderson T. W., Bachadur R. R. Classification into two multivariate normal distribution with different covariance matrices. Annals of mathem stat., 1962, № 2.
26. Cooper D. B. Adaptiv pattern recognition and signal detection using stochastic aprochimation. IEEE Trans, 1964, EC-13, № 3.
27. Cooper D. B., Cooper P. W. Adaptiv pattern recognition and signal detection without supervision. IEEE Int. Conv. Rec., 1964, p. 1.
28. Fralick S. C. Learning to recognize pattern without a teacher. IEEE Trans., 1967, IT-13, № 1.
29. Glaser E. M. Signal detection by adaptive filters. IRE Trans., 1961, IT-7, № 2.
30. Highleyman W. H. Linear dicision function with application to pattern recognition. Proc. IRE, 1962, № 6.
31. Hinich M. J. A model tor a self — adapting filter. Inform. and Control, 1962, Sept.
32. John S. On some classification statistics. Jndian journ. of statistics, 1960, v. 22, p. 3—4.
33. John S. Errors in discrimination. Annals ot mathem. statistics, 1961, № 4.
34. Kabe D. G. Some results on the distribution ot two random matrices used in classification procedures. Annals ot math. statistics, 1963, № 1.
35. Kac M. A note on learning signal detection. IRE Trans., 1962, IT-8, № 2.

36. Kashyap R. L., Blaydon C. C. Recovery of function from noisy measurements taken at randomly selected points and its application to pattern classification. Proc. IEEE, 1966, № 8.
37. Keehn D. G. A note on learning for gaussian properties. IEEE Trans., 1965, IT-11, № 1.
38. Marill T., Green D. Effectiveness of receptors in recognition system. IEEE Trans., 1963, IT-9, № 1.
39. Okatomo M. Discrimination for varience matrices. Annals of mathem. statistics, 1961, № 1—2.
40. Patrick E. A., Hancock J. C. The non — supervised learning of probability spaces and recognition of patterns. IEEE Int. Conv. Rec., 1965, p. VII.
41. Scudder H. J. Adaptive communication receviers. IEEE Trans., 1965, IT-11, № 2.
42. Sebastyen G. Pattern recognition by an adaptive process of sample set construction. IRE Trans., 1962, IT-8, № 3.
43. Spragins J. A note on the iterative application of Bayes rule. IEEE Trans., 1965, IT-11, № 4.
44. Spragins J. Learning without a teacher. IEEE Trans, 1966, IT-12, № 2.
45. Wald A. On a statistical problem arising in the classification of an individual in one of two groups. Annals of mathem. statist., 1944, v. 15.

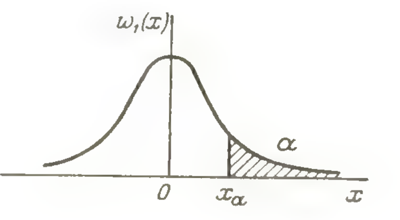
ПРИЛОЖЕНИЯ

Приложение I

Процентные точки нормального распределения

$$w_1(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

α	0,0005	0,0010	0,0015	0,0020	0,0025	0,0030	0,0035	0,0040	0,0045
x_α	3,29053	3,09023	2,96774	2,87816	2,80703	2,74778	2,69684	2,65207	2,61205
α	0,005	0,006	0,007	0,008	0,009	0,010	0,015	0,020	0,025
x_α	2,57583	2,51214	2,45726	2,40892	2,36562	2,32635	2,17009	2,05375	1,95996
α	0,03	0,04	0,05	0,06	0,07	0,08	0,09	0,10	0,12
x_α	1,88079	1,75069	1,64485	1,55477	1,47579	1,40507	1,34076	1,28155	1,17499
α	0,14	0,16	0,18	0,20	0,25	0,30	0,35	0,40	0,45
x_α	1,08032	0,99446	0,91537	0,84162	0,67449	0,52440	0,38532	0,25335	0,12566



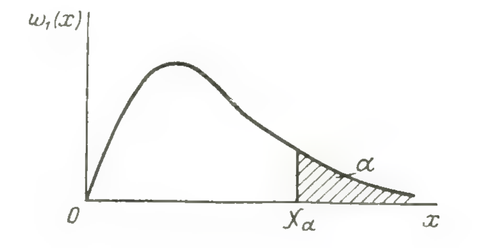
$$x_{0,5+\alpha} = -x_{0,5-\alpha},\quad 0<\alpha<0,5,\quad x_{0,5} = 0.$$

Процентные точки хи-квадрат распределения

$$w_1(x) = \frac{1}{2\Gamma\left(\frac{n}{2}\right)}\left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \; x \geqslant 0.$$

n \ α	0,999	0,995	0,99	0,98	0,975	0,95	0,90	0,80	0,75	0,70	0,50
1	0,05157	0,04393	0,03157	0,03628	0,03982	0,00393	0,0158	0,0642	0,102	0,148	0,455
2	0,00200	0,0100	0,0201	0,0404	0,0506	0,103	0,211	0,446	0,575	0,713	1,386
3	0,0243	0,0717	0,115	0,185	0,216	0,352	0,584	1,005	1,213	1,424	2,366
4	0,0908	0,207	0,297	0,429	0,484	0,711	1,064	1,649	1,923	2,195	3,357
5	0,210	0,412	0,554	0,752	0,831	1,145	1,610	2,343	2,675	3,000	4,351
6	0,381	0,676	0,872	1,134	1,237	1,635	2,204	3,070	3,455	3,828	5,348
7	0,598	0,989	1,239	1,564	1,690	2,167	2,833	3,822	4,255	4,671	6,346
8	0,857	1,344	1,646	2,032	2,180	2,733	3,490	4,594	5,071	5,527	7,344
9	1,152	1,735	2,088	2,532	2,700	3,325	4,168	5,380	6,899	6,393	8,343
10	1,479	2,156	2,558	3,059	3,247	3,940	4,865	6,179	6,737	7,267	9,342
11	1,834	2,603	3,053	3,609	3,816	4,575	5,578	6,989	7,584	8,148	10,341
12	2,214	3,074	3,571	4,178	4,404	5,226	6,304	7,807	8,438	9,034	11,340
13	2,617	3,565	4,107	4,765	5,009	5,892	7,042	8,634	9,299	9,926	12,340
14	3,041	4,075	4,660	5,368	5,629	6,571	7,790	9,467	10,165	10,821	13,339
15	3,483	4,601	5,229	5,985	6,262	7,261	8,547	10,307	11,036	11,721	14,339
16	3,942	5,142	5,812	6,614	6,908	7,962	9,312	11,152	11,912	12,624	15,338
17	4,416	5,697	6,408	7,255	7,564	8,672	10,085	12,002	12,792	13,531	16,338
18	4,905	6,265	7,015	7,906	8,231	9,390	10,865	12,857	13,675	14,440	17,338
19	5,407	6,844	7,633	8,567	8,907	10,117	11,651	13,716	14,562	15,352	18,338
20	5,921	7,434	8,260	9,237	9,591	10,851	12,443	14,578	15,452	16,266	19,337

α / n	0,999	0,995	0,99	0,98	0,975	0,95	0,90	0,80	0,75	0,70	0,50
21	6,447	8,034	8,897	9,915	10,283	11,591	13,240	15,445	16,344	17,182	20,337
22	6,983	8,643	9,542	10,600	10,982	12,338	14,041	16,314	17,240	18,101	21,337
23	7,529	9,260	10,196	11,293	11,688	13,091	14,848	17,187	18,137	19,021	22,337
24	8,085	9,886	10,856	11,992	12,401	13,848	15,659	18,062	19,037	19,943	23,337
25	8,649	10,520	11,524	12,697	13,120	14,611	16,473	18,940	19,939	20,867	24,337
26	9,222	11,160	12,198	13,409	13,844	15,379	17,292	19,820	20,843	21,792	25,336
27	9,803	11,808	12,879	14,125	14,573	16,151	18,114	20,703	21,749	22,719	26,336
28	10,391	12,461	13,565	14,847	15,308	16,928	18,939	21,588	22,657	23,647	27,336
29	10,986	13,121	14,256	15,574	16,047	17,708	19,768	22,475	23,567	24,577	28,336
30	11,588	13,787	14,953	16,306	16,791	18,493	20,599	23,364	24,478	25,508	29,336



α / n	0,30	0,25	0,20	0,10	0,05	0,025	0,02	0,01	0,005	0,001
1	1,074	1,323	1,642	2,706	3,841	5,024	5,412	6,635	7,879	10,827
2	2,408	2,773	3,219	4,605	5,991	7,378	7,824	9,210	10,597	13,815
3	3,665	4,108	4,642	6,251	7,815	9,348	9,837	11,345	12,838	16,268
4	4,878	5,385	5,989	7,779	9,488	11,143	11,668	13,277	14,860	18,465
5	6,064	6,626	7,289	9,236	11,070	12,832	13,388	15,086	16,750	20,517
6	7,231	7,841	8,558	10,645	12,592	14,449	15,033	16,812	18,548	22,457
7	8,383	9,037	9,803	12,017	14,067	16,013	16,622	18,475	20,278	24,322
8	9,524	10,219	11,030	13,362	15,507	17,535	18,168	20,090	21,955	26,125
9	10,656	11,389	12,242	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	11,781	12,549	13,442	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	12,899	13,701	14,631	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	14,011	14,845	15,812	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	15,119	15,984	16,985	19,812	22,362	24,736	25,472	27,688	29,819	34,528
14	16,222	17,117	18,151	21,064	23,685	26,119	26,873	29,141	31,319	36,123
15	17,322	18,245	19,311	22,307	24,996	27,488	28,259	30,578	32,801	37,697
16	18,418	19,369	20,465	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	19,511	20,489	21,615	24,769	27,587	30,191	30,995	33,409	35,718	40,790
18	20,601	21,605	22,760	25,989	28,869	31,526	32,346	34,805	37,156	42,312

Продолжение прилож. II

α / n	0,30	0,25	0,20	0,10	0,05	0,025	0,02	0,01	0,005	0,001
19	21,689	22,718	23,900	27,204	30,144	32,852	33,687	36,191	38,582	43,820
20	22,775	23,828	25,038	28,412	31,410	34,170	35,020	37,566	39,997	45,315
21	23,858	24,935	26,171	29,615	32,671	35,479	36,343	38,932	41,401	46,797
22	24,939	26,039	27,301	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	26,018	27,141	28,429	32,007	35,172	38,076	38,968	41,638	44,181	49,728
24	27,096	28,241	29,553	33,196	36,415	39,364	40,270	42,980	45,558	51,179
25	28,172	29,339	30,675	34,382	37,652	40,646	41,566	44,314	46,928	52,620
26	29,246	30,434	31,795	35,563	38,885	41,923	42,856	45,642	48,290	54,052
27	30,319	31,528	32,912	36,741	40,113	43,194	44,140	46,963	49,645	55,476
28	31,391	32,620	34,027	37,916	41,337	44,461	45,419	48,278	50,993	56,893
29	32,461	33,711	35,139	39,087	42,557	45,722	46,693	49,588	52,336	58,302
30	33,530	34,800	36,250	40,256	43,773	46,979	47,962	50.892	53,672	59,703

Примечание. При $n \gg 1$

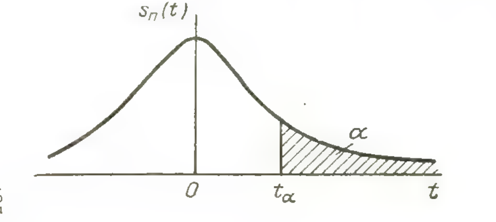
$$\chi^2_\alpha \sim \frac{1}{2}(x_\alpha \pm \sqrt{2n-1})^2,$$

где x_α — процентная точка нормального распределения (см. приложение I).

ПРОЦЕНТНЫЕ ТОЧКИ РАСПРЕДЕЛЕНИЯ СТЬЮДЕНТА

Значения t_α, удовлетворяющие равенству $\int\limits_{t_\alpha}^{\infty} s_n(t)\,dt = \alpha$

$n-1$ \ α	0,45	0,4	0,35	0,3	0,25	0,2	0,15	0,1	0,05	0,025	0,01	0,005	0,0005
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	142	289	445	617	0,816	1,061	1,336	1,886	2,920	4,303	6,965	9,965	31,598
3	137	277	424	584	765	0,978	1,250	1,638	2,353	3,181	4,541	5,841	12,941
4	134	271	414	569	741	941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	132	267	408	559	727	920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	131	265	404	553	718	906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	130	263	402	549	711	896	1,119	1,415	1,895	2,365	2,998	3,499	5,405
8	130	262	399	546	706	889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	129	261	398	543	703	883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	129	260	397	542	700	879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
12	128	259	395	539	695	873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
14	128	258	393	537	692	868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
16	128	258	392	535	690	865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
18	127	257	393	534	688	862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
20	127	257	391	533	687	860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
25	127	256	390	531	684	856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
30	126	256	389	530	683	854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
60	126	254	387	527	679	848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
∞	126	253	385	524	674	842	1,036	1,282	1,645	1,960	2,326	2,576	3,291



$0{,}5+\alpha = -t_{0,5-\alpha},\ 0<\alpha<0{,}5,\ t_{0,5}=0.$

Корреляционный эллипсоид

Рассмотрим совокупность n случайных величин $\xi_1, \ldots, \xi_n$. Обозначим среднее значение случайной величины ξ_i через a_i, а ковариацию случайных величин ξ_i и ξ_j — через $r_{ij} = m_1\{(\xi_i - a_i)(\xi_j - a_j)\}$. Матрица $\mathbf{M}$, составленная из элементов r_{ij} $(i, j = 1, \ldots, n)$, представляет корреляционную матрицу совокупности случайных величин.

Эллипсоид, уравнение которого имеет вид

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{D_{ij}}{D}(z_i - a_i)(z_j - a_j) = 1, \quad D > 0, \tag{1}$$

называется *корреляционным*. В уравнении (1) D_{ij} — алгебраические дополнения элемента a_{ij} в матрице $\mathbf{M}$, а D — ее детерминант. Линейным (ортогональным) преобразованием переменных квадратичная форма в левой части (1) приводится к сумме квадратов

$$\sum_{k=1}^{n}\frac{y_k^2}{\lambda_k} = 1. \tag{2}$$

Здесь λ_k — характеристические числа корреляционной матрицы, являющиеся корнями уравнения

$$|\mathbf{M} - \lambda\mathbf{I}| = 0,$$

где $\mathbf{I}$ — единичная матрица.

Уравнение (2) представляет n-мерный эллипсоид с полуосями, равными $\sqrt{\lambda_k}$, к которому приводится эллипсоид (1) вращением (ортогональным преобразованием). Объем рассматриваемого эллипсоида равен

$$V = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}\sqrt{D}. \tag{3}$$

Таким образом, квадрат объема эллипсоида пропорционален величине детерминанта D, которую иногда называют *обобщенной дисперсией*. При $n = 1$ эта величина равна

$$D = m_1\{(\xi - a)^2\} = \sigma^2,$$

т. е. совпадает с обычной дисперсией случайной величины, а при $n=2$

$$D=\begin{vmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{vmatrix}=\sigma_1^2\sigma_2^2(1-r),$$

где σ_1^2, σ_2^2 и r — дисперсии и коэффициент корреляции двух случайных величин.

Если корреляционный эллипсоид (1) растянуть в $\sqrt{n+2}$ раз, то получаем *эллипсоид рассеяния*. Для совокупности нормальных случайных величин плотность вероятности на эллипсоиде рассеяния постоянна (см. первую книгу, стр. 59 и 66). Для произвольного распределения корреляционный эллипсоид обладает тем свойством, что совокупность случайных величин, распределенных *равномерно* по области n-мерного пространства, ограниченного этим эллипсоидом, имеет те же первые и вторые моменты (ковариации), как и заданная совокупность случайных величин.

Приложение V

Регрессия

Пусть ξ и η — зависимые случайные величины, характеризуемые совместной плотностью вероятностей $w_2(x, y)$. Условное среднее значение η при $\xi = x$ (см. (2.108) в первой книге), рассматриваемое как функция переменного x, задает *кривую регрессии*

$$z=m_1\{\eta\,|\,x\}=\frac{\int\limits_{-\infty}^{\infty} y w_2(x, y)\,dy}{\int\limits_{-\infty}^{\infty} w_2(x, y)\,dy}=\int\limits_{-\infty}^{\infty} y w(y\,|\,x)\,dy. \qquad (1)$$

Нетрудно доказать (см. для аналогии задачу 2.6 в первой книге), что среди всех возможных функций $g(x)$ минимум среднеквадратического отклонения $m_1\{[\eta-g(\xi)]^2\}$ достигается при

$$g(x)=m_1\{\eta\,|\,x\}. \qquad (2)$$

В некоторых случаях рассматривают аппроксимацию η посредством ξ по критерию минимума среднего квадрата отклонения η от $f(\xi)$ для заданного класса функций $g(x)$, определенных с точностью до неизвестных параметров.

Например, если $g(x) = a_1 x + a_2$, то ищут такие значения a_1, a_2, для которых $m_1\{(\eta - a_1\xi - a_2)^2\}$ минимальны. Прямая $z = a_1 x + a_2$ в этом случае называется прямой средней квадратической регрессии. Более общим является случай полиномиальной средней квадратической регрессии, когда

$$g(x) = a_1 x^n + a_2 x^{n-1} + \ldots + a_{n+1}.$$

В случае линейной средней квадратической регрессии минимум $m_1\{(\eta - a_1\xi - a_2)^2\}$ достигается при

$$a_1 = r\sqrt{\frac{M_2\{\eta\}}{M_2\{\xi\}}}, \quad a_2 = m_1\{\eta\} - a_1 m_1\{\xi\}, \qquad (3)$$

где r — коэффициент корреляции случайных величин ξ и η. Прямая средней квадратической регрессии при этом задается уравнением

$$z = m_1\{\eta\} + [x - m_1\{\xi\}]\frac{M_2\{\xi\eta\}}{M_2\{\xi\}}. \qquad (4)$$

Если ξ и η — зависимые нормальные случайные величины, то правая часть (4) в точности совпадает с условным средним η при $\xi = x$ (см. стр. 87 в первой книге). Отсюда следует, что прямая средней квадратической регрессии для нормальных случайных величин совпадает с кривой регрессии. Иначе говоря, линейная аппроксимация η посредством ξ по критерию минимума среднего квадрата ошибки является наилучшей.

Приведенные понятия обобщаются на произвольную конечную совокупность взаимозависимых случайных величин $\xi_1, \ldots, \xi_n$. Условное среднее значение ξ_1 при $\xi_i = x_i$, $i = 2, \ldots, n$, равно

$$m_1\{\xi_1 \mid x_2, \ldots, x_n\} = \frac{\int\limits_{-\infty}^{\infty} x_1 w_n(x_1, \ldots, x_n)\,dx_1}{\int\limits_{-\infty}^{\infty} w_n(x_1, \ldots, x_n)\,dx_1}. \qquad (5)$$

Геометрическое место точек $(m_1\{.\}, x_2, \ldots, x_n)$ при всевозможных значениях $x_2, \ldots, x_n$ есть *гиперповерхность регрессии*. Может быть определена *гиперплоскость средней квадратической регрессии*

$$z_1 = m_1\{\xi_1\} + \sum_{i=2}^{n} \rho_{1i}\,[x_i - m_1\{\xi_i\}]. \qquad (6)$$

Величины ρ_{1i} определяются из системы линейных уравнений

$$\sum_{j=2}^{n} b_{ij}\rho_{1j} = b_{i1}, \qquad i = 2, \ldots, n, \tag{7}$$

где

$$b_{ij} = m_1\{\xi_i \xi_j\}, \tag{7'}$$

так что, обозначая через D_{ij} алгебраические дополнения в матрице $\|b_{ij}\|$, имеем

$$\rho_{1i} = \frac{D_{1i}}{D_{11}}. \tag{8}$$

Можно показать, что гиперповерхность регрессии для совокупности *нормальных* случайных величин совпадает с гиперплоскостью средней квадратической регрессии.

32 - 197

НАИБОЛЕЕ УПОТРЕБИТЕЛЬНЫЕ ОБОЗНАЧЕНИЯ

A, A_0 — амплитуда сигнала,
$A(\omega)$ — амплитудный спектр,
$A(t)$ — квадратурная составляющая процесса,
a — среднее значение нормальной случайной величины, амплитуда детерминированного сигнала,
$B(t, y)$ — корреляционная функция случайного процесса,
$B(\tau)$ — корреляционная функция стационарного в широком смысле процесса,
$B_{\xi\eta}(t, y)$ — взаимная корреляционная функция случайных процессов $\xi(t)$ и $\eta(t)$,
$B_{\xi\eta}(\tau)$ — взаимная корреляционная функция стационарно связанных в широком смысле случайных процессов $\xi(t)$ и $\eta(t)$,
b_n — смещение оценки,
$C(t)$ — квадратурная составляющая процесса,
c, c^* — пороги,
$C(\omega)$ — частотная характеристика линейной системы,
D — детерминант корреляционной матрицы,
d_N, d_T — параметры рабочих характеристик обнаружения сигнала,
$E(t)$ — огибающая случайного процесса,
E — энергия сигнала,
e — относительная эффективность оценки,
$F_N(x_1, \ldots, x_N)$ — N-мерная интегральная функция распределения,
$F(x)$ — интеграл Лапласа,
$F(\omega)$ — энергетический спектр стационарного (в широком смысле) случайного процесса,
$F_{\xi\eta}(\omega)$ — взаимный энергетический спектр ста-

ционарно связанных (в широком смысле) случайных процессов $\xi(t)$ и $\eta(t)$,

$f\{\Pi\}$ — критерий качества,
G — область пространства выборок,
H — гипотеза,
$H(s)$ — энтропия,
$H(s|\gamma)$ — условная энтропия,
$h(t, y, u)$ — резольвента,
$h(u, v)$ — импульсная переходная функция линейной системы, весовая функция
$h^*(u, v)$ — импульсная переходная функция оптимальной линейной системы,
I_n — информация по Фишеру,
I — количество информации,
J_ν — бесселева функция 1-го рода ν-го порядка,
I_ν — бесселева функция 1-го рода ν-го порядка от чисто мнимого аргумента,
$K(u, v)$ — ядро квадратичной интегральной формы, характеристики фильтра второго порядка,
$k^*(i\omega)$ — передаточная функция оптимальной линейной системы,
$k(i\omega)$ — передаточная функция линейной системы,
$\mathrm{k}(z)$ — функция распределения Колмогорова,
K — порог,
L — функция правдоподобия,
l — отношение правдоподобия,
m^2 — второй момент случайной амплитуды сигнала,
M_k — k-й центральный момент,
M_2 — дисперсия,
M_2^* — выборочная дисперсия,
m_k — k-й момент,
m_1 — среднее,
m_1^* — выборочное среднее,
m — число состояний, число непересекающихся областей пространства выборок,

N — размерность многомерного априорного распределения,
N_0 — спектральная плотность белого шума,
n — размер выборки,
P — вероятность события,
p — априорная вероятность,
Π — функция потерь,
$Q_n(x)$ — полином из совокупности ортогональных полиномов,
q — априорная вероятность,
$r(t)$ — реализация огибающей,
$r(\vartheta)$ — условная функция риска,
R — средний риск,
R^* — байесовский риск,
$R(\tau)$ — коэффициент корреляции,
S — пространство параметров,
$s(t)$ — детерминированный процесс (сигнал),
s — параметр распределения,
s_k — возможные состояния,
$s_n(x)$ — плотность вероятности Стьюдента,
T — время наблюдения,
t — текущее время, нормированная ошибка оценки,
U — число инверсий,
$u(t)$ — квадратурная составляющая детерминированного сигнала,
$V(t)$ — решение неоднородного линейного интегрального уравнения,
$v(t)$ — квадратурная составляющая детерминированного сигнала,
$\mathbf{W}_1(x)$, $w_1(x)$, $W_1(x)$ — одномерные функции распределения случайной величины,
$w_N(x_1, \ldots, x_N)$ — многомерная функция распределения совокупности случайных величин,
$W_N(x_1, \ldots, x_N)$ — функция правдоподобия выборки размером n,
$\mathbf{X}$ — вектор выборочных значений,
x_α — процентное отклонение случайной величины,
x_i — выборочное значение,

$x(t)$ — реализация случайного процесса,
$z(t)$ — комплексная огибающая,
$Z_T(i\omega)$ — спектр усеченной реализации случайного процесса,
z — комплексная переменная,
$\bar{z}$ — величина, комплексно сопряженная с z,
α — условная вероятность ошибки первого рода, условная вероятность ложной тревоги,
β — условная вероятность ошибки второго рода, условная вероятность пропуска сигнала,
$\Gamma(x)$ — полная гамма-функция,
$\Gamma(x, y)$ — неполная гамма-функция,
γ — решение; коэффициент доверия,
$\delta(x)$ — дельта-функция,
$\delta(\gamma \mid \mathbf{x})$ — правило выбора решения,
Δ — ширина полосы частот,
ε — относительная длина доверительного интервала,
ε^2 — среднеквадратическая ошибка,
Λ — усредненная функция правдоподобия,
λ_k — собственное число линейного интегрального уравнения,
μ — отношение априорных вероятностей; параметр марковского процесса,
Ψ — функция неопределенности,
$\psi(t)$ — реализация фазы,
ξ — случайная величина,
$\xi(t)$ — случайный процесс,
η — случайная величина,
$\eta(t)$ — случайный процесс,
χ^2 — случайная величина, распределенная по закону хи-квадрат,
Im — мнимая часть,
Re — действительная часть,
$\binom{n}{k}$ — число сочетаний из n элементов по k,
$\in$ — символ, обозначающий принадлежность множеству.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Н

О

С

Т

У

Ф

Ч

Э

ОГЛАВЛЕНИЕ

БОРИС РУВИМОВИЧ ЛЕВИН

ТЕОРЕТИЧЕСКИЕ ОСНОВЫ СТАТИСТИЧЕСКОЙ РАДИОТЕХНИКИ

Редактор А. А. Александрова
Художественный редактор В. Т. Сидоренко
Технический редактор З. Н. Ратникова
Корректоры Л. И. Кирильченко, Л. Г. Кучина

Сдано в набор 9/I—68 г. Подписано к печати 18/VI—68 г.
Т-08464 Формат 84 × 108/32 Бумага типографская № 1
Объем 26,46 п. л. Уч.-изд. л. 26,321. Тираж 35 000 экз.
Издательство «Советское радио», Москва, Главпочтамт, п/я 693
Зак. 197.

Московская типография № 16 Главполиграфпрома
Комитета по печати при Совете Министров СССР.
Трехпрудный пер., 9.
Цена в переплете № 7 — 1 р. 96 к.